Network Science Summer School

net-science.github.io

Universiteit Utrecht

Instructors

Vincent Buskens
Sociology
Utrecht University
(not in 2025)

Jiamin Ou
Sociology
Utrecht University

Leto Peel
Data Analytics & Digitalisation
Maastricht University

Javier Garcia-Bernardo
Methodology & Statistics
Utrecht University

Mahdi Shafiee Kamalabad
Methodology & Statistics
Utrecht University

Elena Candellone
Methodology & Statistics
Utrecht University

Participants

Why are you taking this summer school?

Program

Monday:

Introduction to networks (Javier)

Network centrality (Javier)

Tuesday:

Network models (Leto)

Statistical models (Mahdi)

Wednesday :

Community detection (Leto)

Link prediction (Javier & Leto)

Thursday:

Network inference (Mahdi)

Friday:

Dynamics in networks (Jiamin)

Day program

9:00–12:00:

Introduction to network science

12:00-13:00

Lunch

13:00–16:30:

Network representation

Centrality

Network game

Introduce yourself, and find one thing you have in common:

- Countries (apart from the NLD) that you have lived in
- Favorite cuisine
- Sports you practice
- Programming languages you use
- …

Draw a line in the whiteboard, write the names in this spreadsheet:
https://tinyurl.com/network-game

It's a public file, please don't use your full name

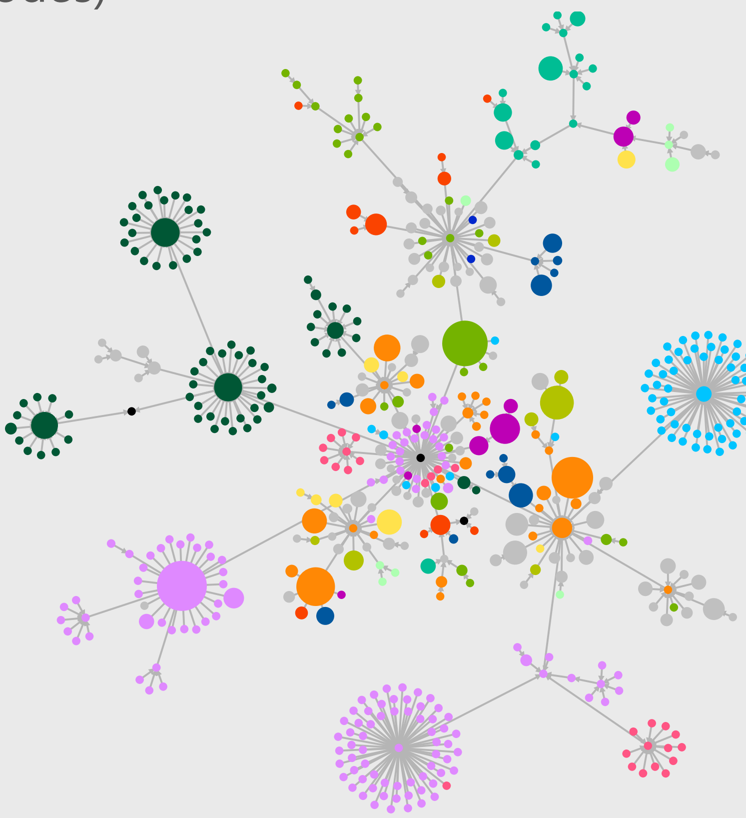
Introduction to networks

What is a network?

Mathematical representation of relationships (edges) between entities (nodes)

The most important question to ask yourself:

What are the nodes, and what are the edges?



Types of networks

	Network	Nodes	Edges
Social/ Behavioral	Friendship	People	Friendships
	Follower	Online accounts	Followers/likes
	Psychological	Symptoms	Co-ocurrence
Biology	Gene regularory	Genes	Activations/inhibations
	Food web	Animals	Predating
Economic	Trade	Countries/companies	Money flows
	Ownership	Companies	Ownership stakes
	Stock correlations	Listed companies	Co-movements
Intrastructure	Internet	Computers (IPs)	Data transmission
	Power grid	Power stations	Power lines
	Airplane network	Airports	Flights

https://aaronclauset.github.io/courses/5352/csci5352_F21_L1.pdf

Type of networks and characteristics

1: Interaction and flow → "Real networks".

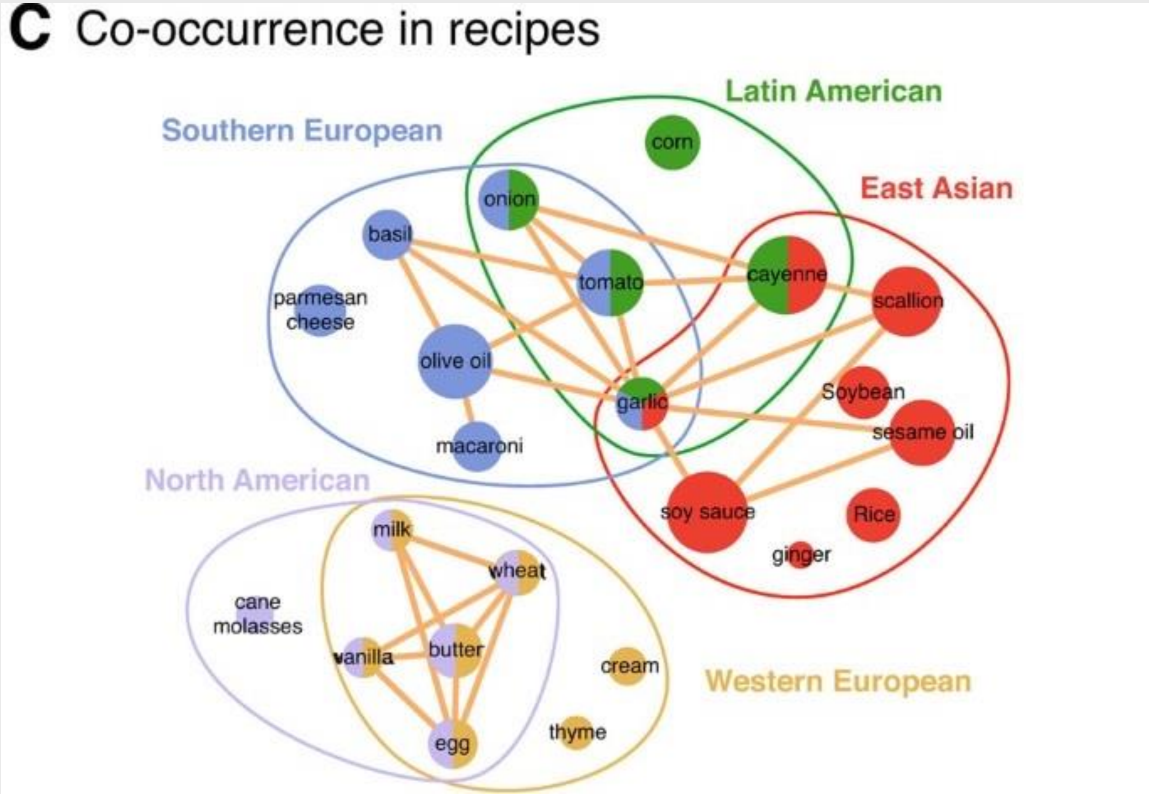
- Offline interactions
- Online interactions

2: Affiliation → Node 1 is part of/related to node 2

- e.g. students in classrooms
- Bipartite networks

3: Co-occurrence → Node 1 is correlated with node 2

- Stock market networks
- Brain networks



Relational data (networks)

Our life is completely defined by networks: relationships, interactions, communications.

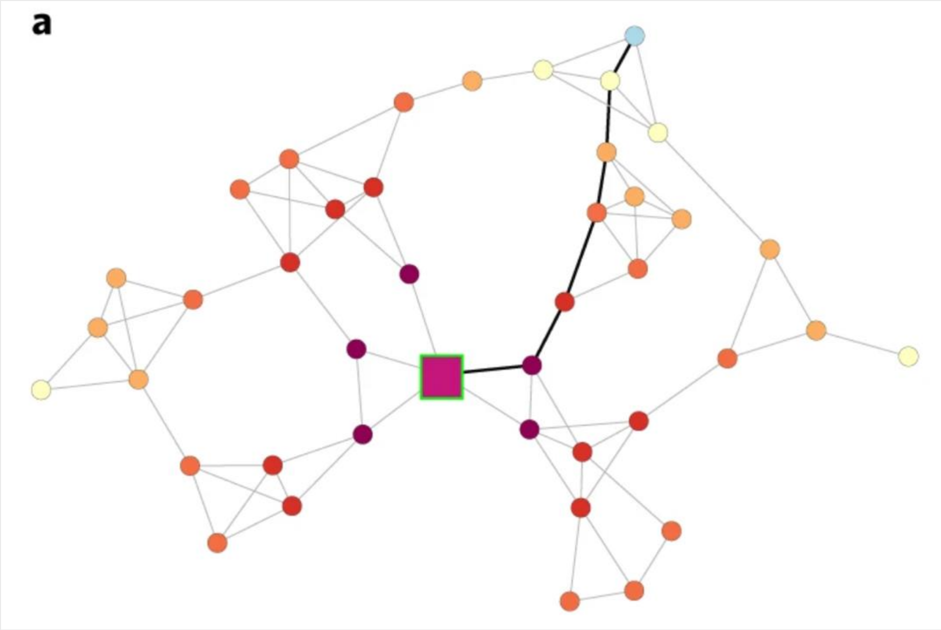
***Biological networks** governing the interactions between genes in our cells determine our development, **neural networks** in our brain make us think, **information networks** guide our knowledge and culture, **transportation networks** allow us to move, **and social networks** sustain our life.*

A First Course in Network Science, F Menczer, S Fortunato, C.A. Davis

Tabular data

	outcome	Age	Income	...
Person 1	1	30	50000	
Person 2	0	25	30000	
Person 3	1	32	40000	
Person 4	0	40	25000	
Person 5	0	28	100000	

Network data



How do we study the connections?

Why do we care about the connections?

1) They reflect underlying patterns (e.g. differences in power/preferences/roles/groups).

Network of countries trading with each other: Which country has the most bargaining power?

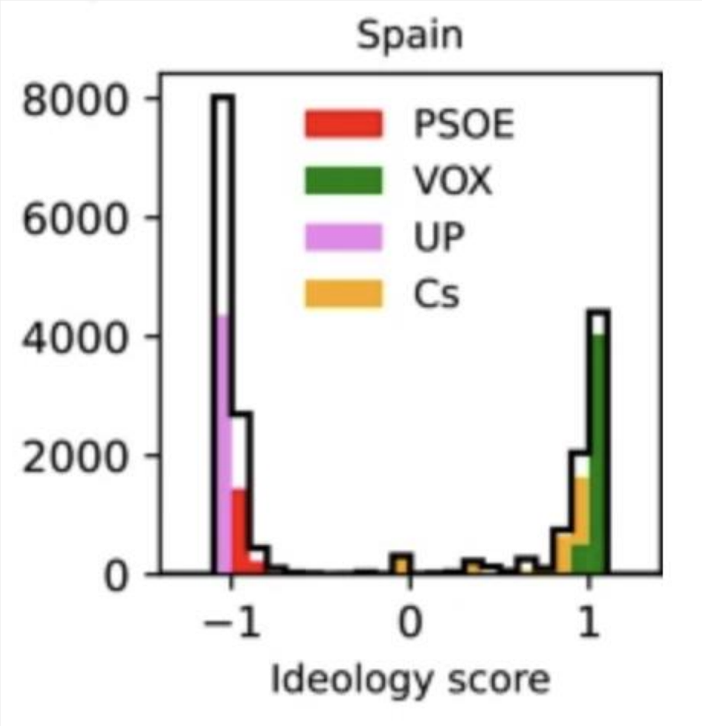
Network of politicians and their votes to bills: Is the political system polarized?

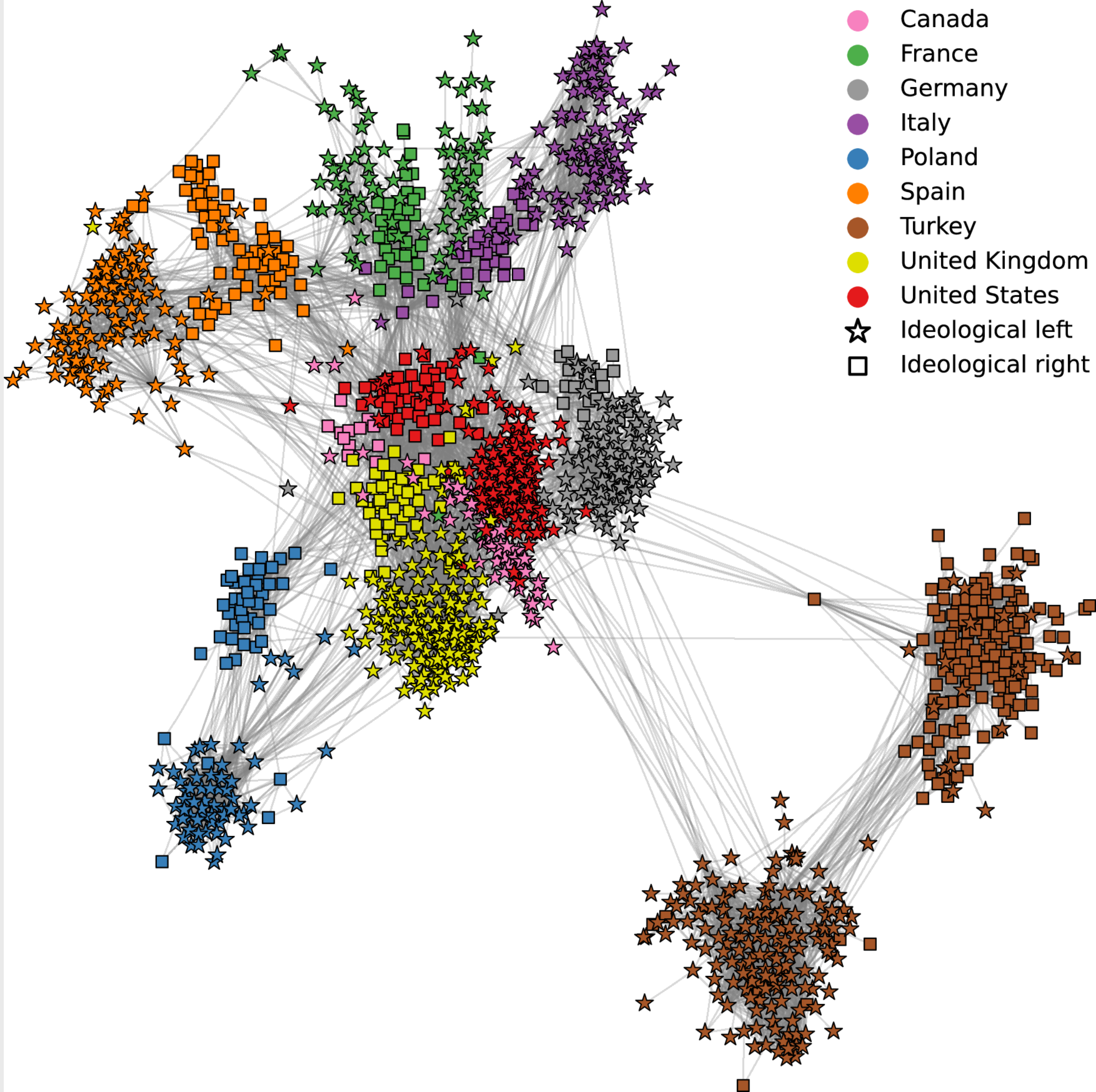
Networks of genes regulating each other: Which combination of genes regulate specific biological functions?

Network of friends: How does social capital help your job opportunities?

Food webs: Why do some ecosystems collapse?

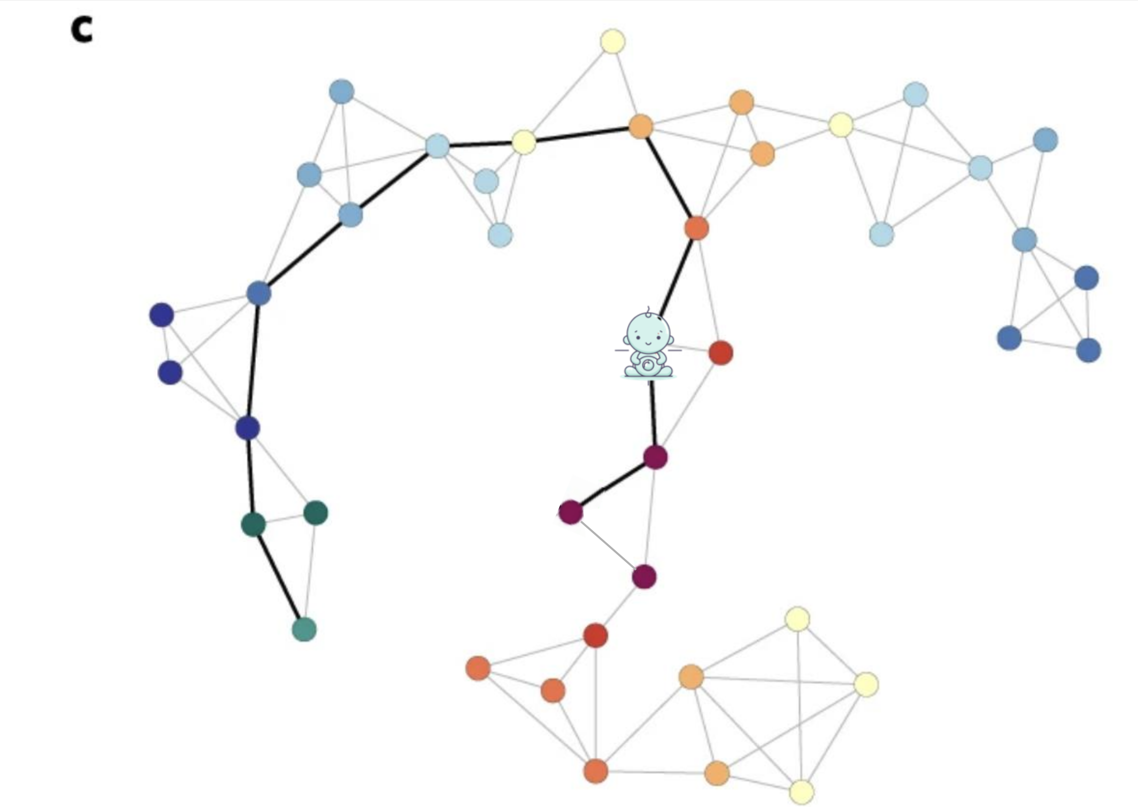
Today (Javier), Tuesday and Wednesday (Leto)





Falkenberg et al. (2024). *Nature Communications*

2) They constrain/facilitate future change



Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. Nature human behaviour, 4(6), 588-596.

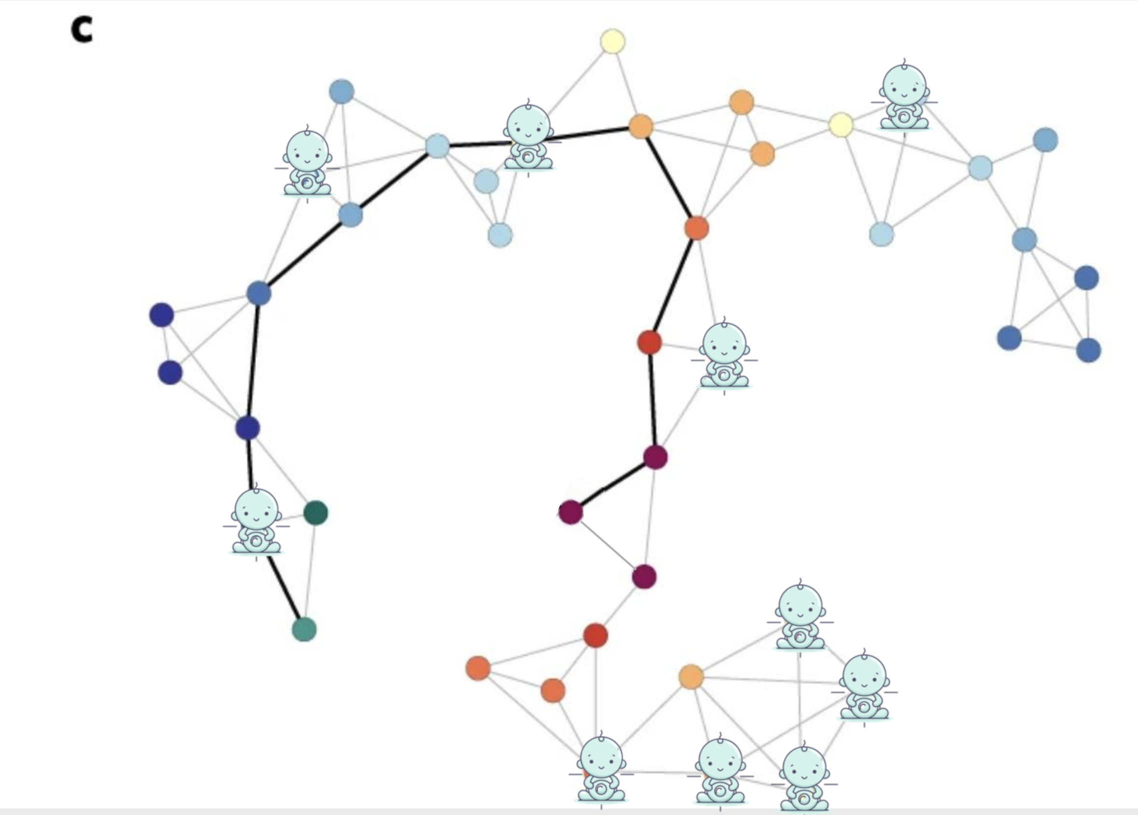
Which person would you vaccinate first?

app.wooclap/NETSCI

Networks allow us to understand dynamics: epidemics, contagion, development

Tuesday (Mahdi) and Friday (Vincent and Jiamin)

2) They constrain/facilitate future change



Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. Nature human behaviour, 4(6), 588-596.

Which person will become a parent next?

app.wooclap/NETSCI

Tuesday (Mahdi) and Friday (Vincent and Jiamin)

3) They can create *emergence*

Sometimes a system (e.g. a society) has properties that the individual parts do not have.
These properties are called "*emergent*" properties, and the system a *complex* system.

"There's no love in a carbon atom, no hurricane in a water molecule,

no financial collapse in a dollar bill." *Peter Sheridan Dodds*

In social science: Connect micro-behavior to macro-outcomes.

e.g. Schelling model: why do we see urban segregation?

Every actor lives in a house and is connected to its neighbors in a network. Every actor is the same:

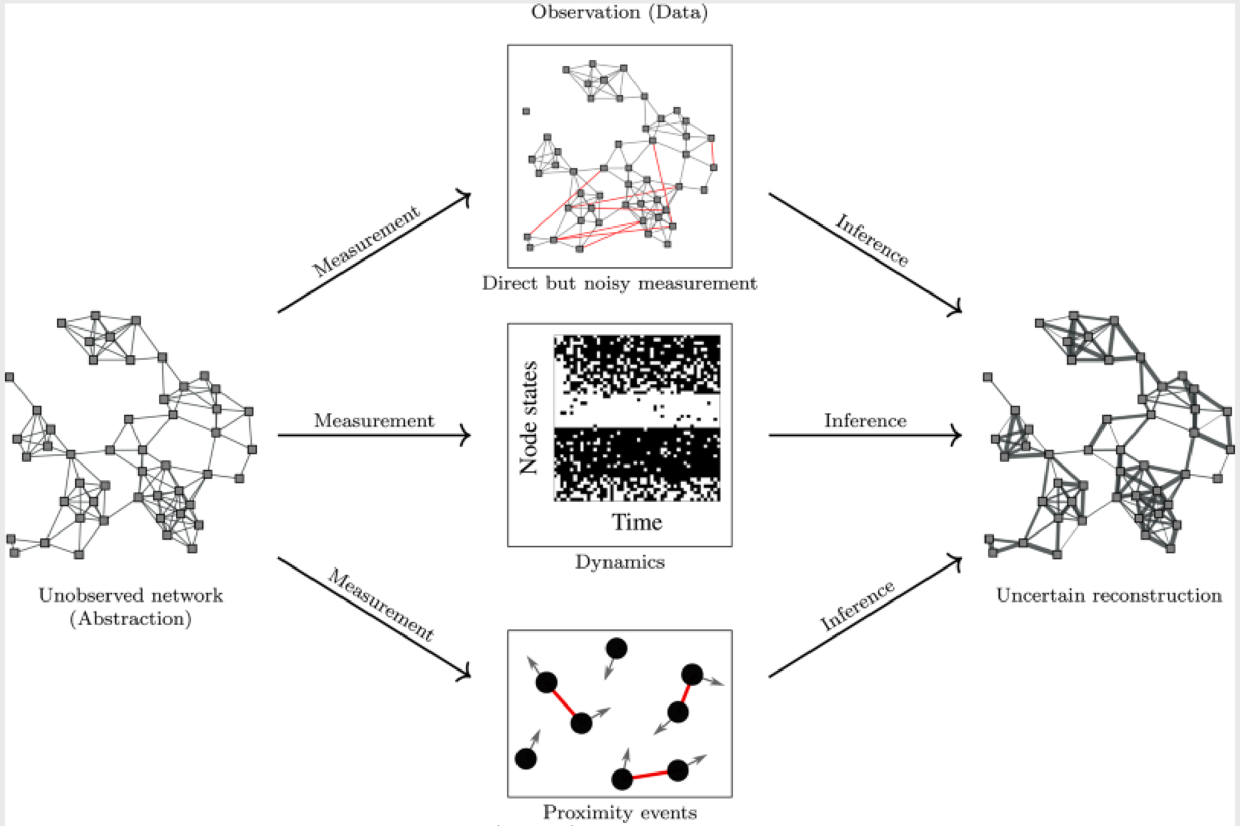
- They want to have 1/3 of their neighbors to be like them
- Otherwise, they move to a random house

http://nifty.stanford.edu/2014/mccown-schelling-model-segregation/

4) They are unknown

You may have a noisy observations and would like to understand who is connected to whom.

- Gene regulatory networks: Targets for new drugs
- Financial networks: Avoid unknown indirect risks
- Social networks: People are extremely unreliable



Peel, Peixoto and De Domenico (2022)

Comparing networks:

The network of a high-school in French collected with four ways:

- Proximity sensors
- Facebook friends
- Diaries (keeping track of their contacts)
- Surveys (recollecting their contacts)

http://javier.science/panel_network/

Tuesday (Leto), Wednesday (Javier & Leto) and Thursday (Mahdi)

Examples of research questions that can be answered with networks

Understanding Complex Systems: Find insights that we would miss if we would study the nodes independently (one person != society)

Epidemiology: How to stop disease transmission in a social network?

Criminology: How to detect criminal actors in a network of money flows?

Biotechnology: Which genes to target to stop cancer in a gene regulatory network?

Ecology: Which animals we need to preserve to avoid ecosystem collapse?

Psychology: How does attitude change depend on the correlation between attitudes?

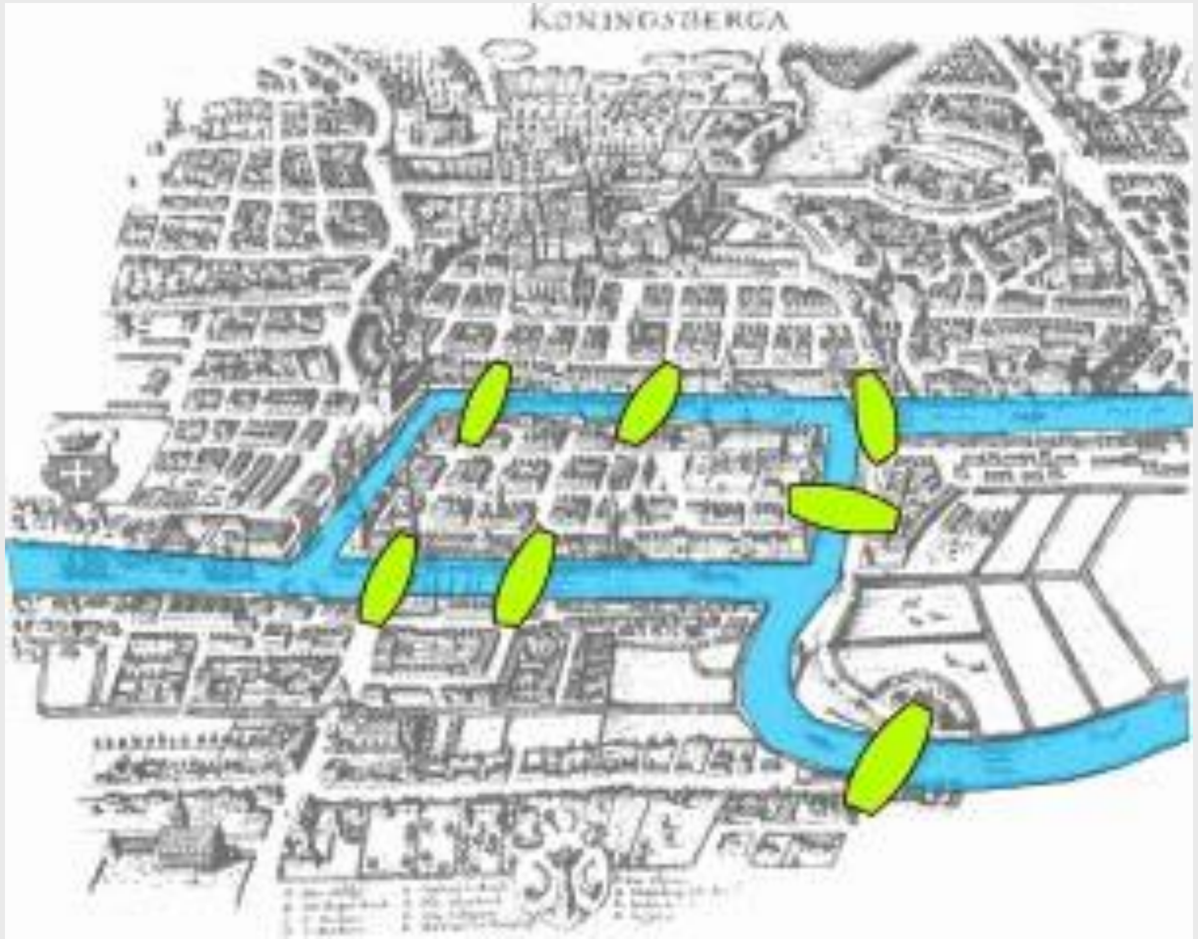
Engineering: How to improve network performance and reliability in power grids?

Economics: How does country development depend on the type of products a country export?

Social science: How does social capital affect upward mobility?

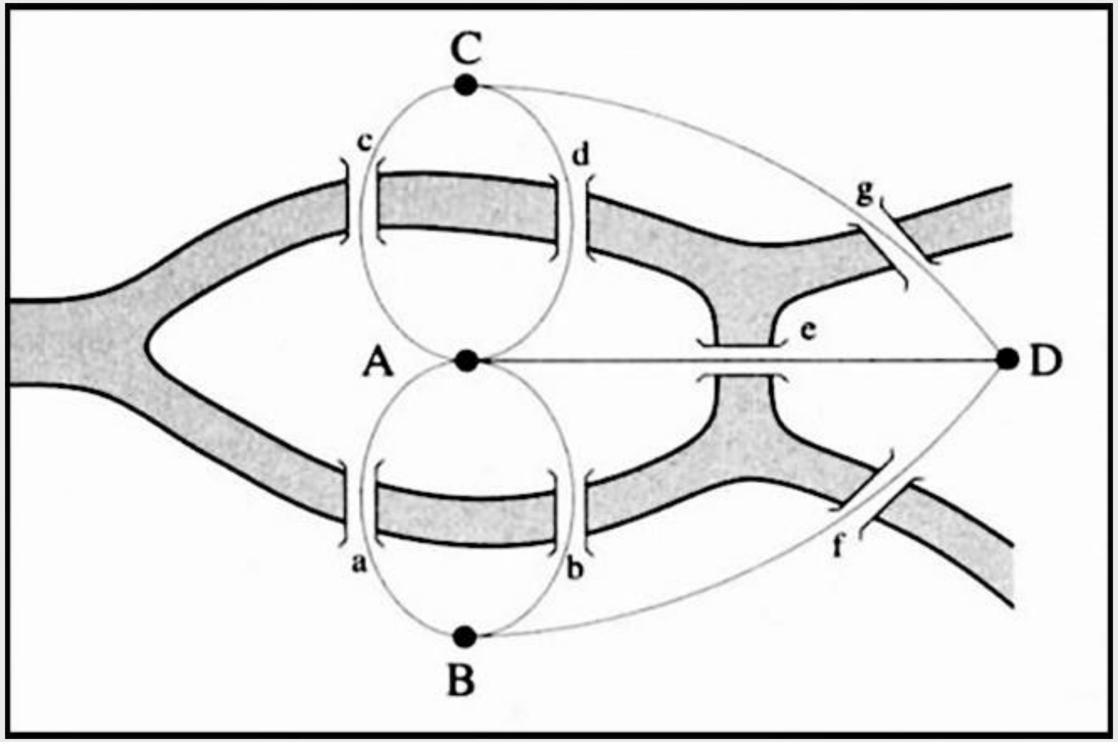
Physics view: Dependence on topology (reliability, dynamics, emergent behavior and phase transitions)

Tiny bit of history: Bridges of Königsberg

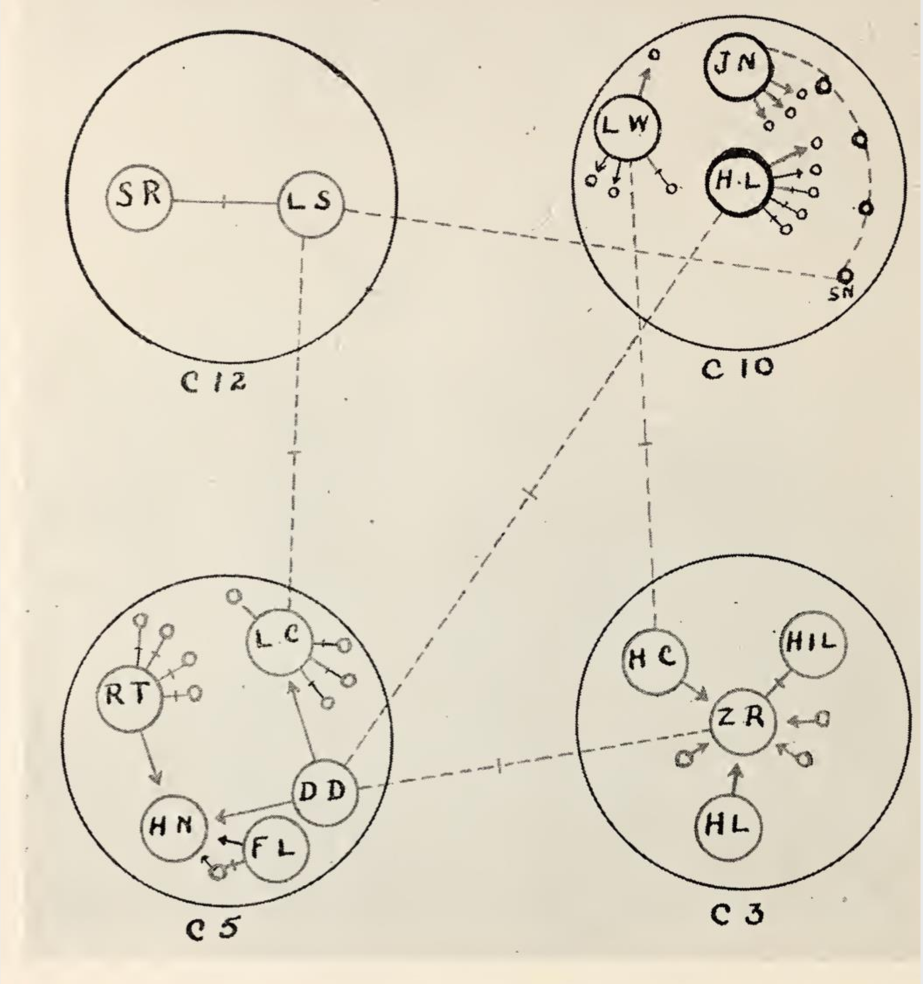


Is there a way to cross each bridge exactly once and return to the starting point

Euler (1736)



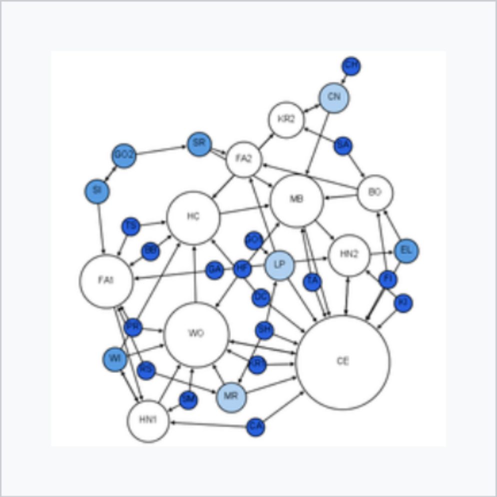
Network Science has roots in sociology



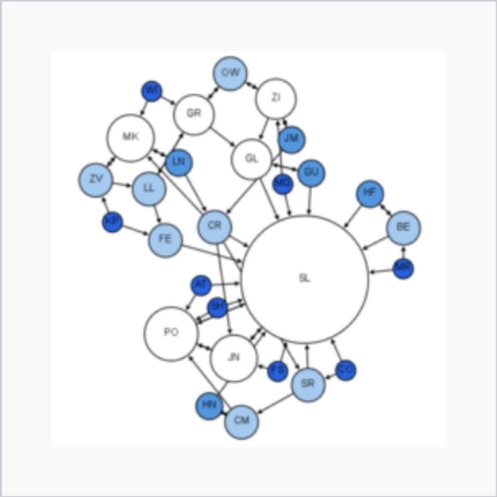
Moreno. Who shall survive?

Network science: Social and behavioral scientists in the XX century (e.g. Jacob Moreno & Hellen Hall Jennings, Harrison White, Mark Granovetter)

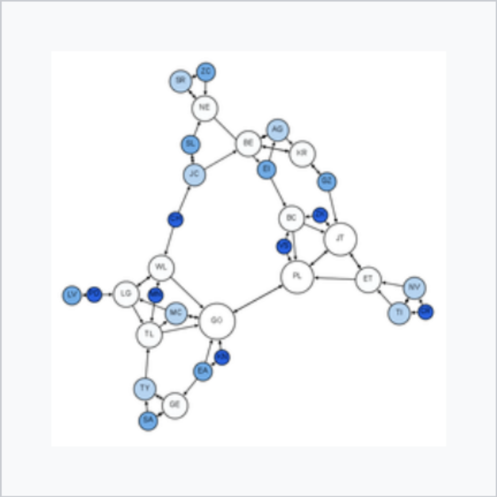
- Hellen Hall Jennings and Jacob Moreno (1930s): Hudson School for girls: Sociometry. The location of girls in networks determined whether and when they would run away.
- Manfred Kochen & Ithiel de Sola Pola (1950s), Jeffrey Travers and Stanley Milgram's (1969): Small-world studies
- Nancy Howell (1969): *The Search for an Abortionist*, women acquired scarce information through short chains of weak ties.
- Mark Granovetter (1973) *The Strength of Weak Ties*. Diffusion of information takes place primarily through bridges (weak ties). Strong links are redundant.
- François Lorrain & Harrison White (1976): Blockmodels for networks
- Duncan Watts, Steven Strogatz (1998): Next wave of network science



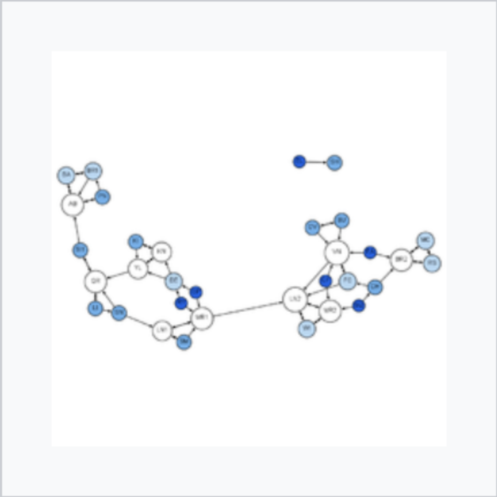
1st Grade



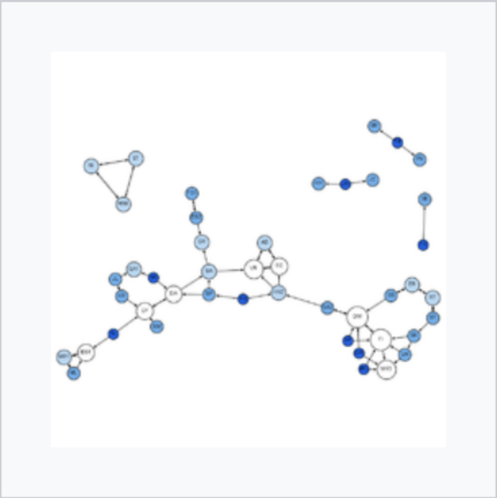
2nd Grade



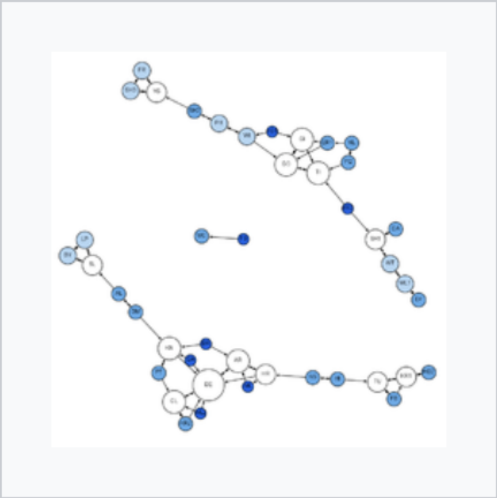
3rd Grade



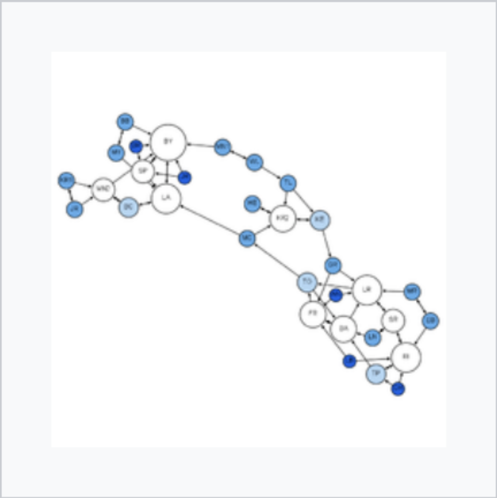
4th Grade



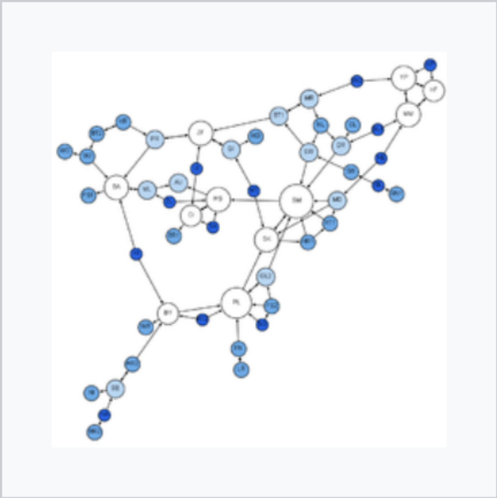
5th Grade



6th Grade



7th Grade



8th Grade

Moreno, source: wikipedia. Edges: who wants to be sitting next to each other?

Three key concepts of today

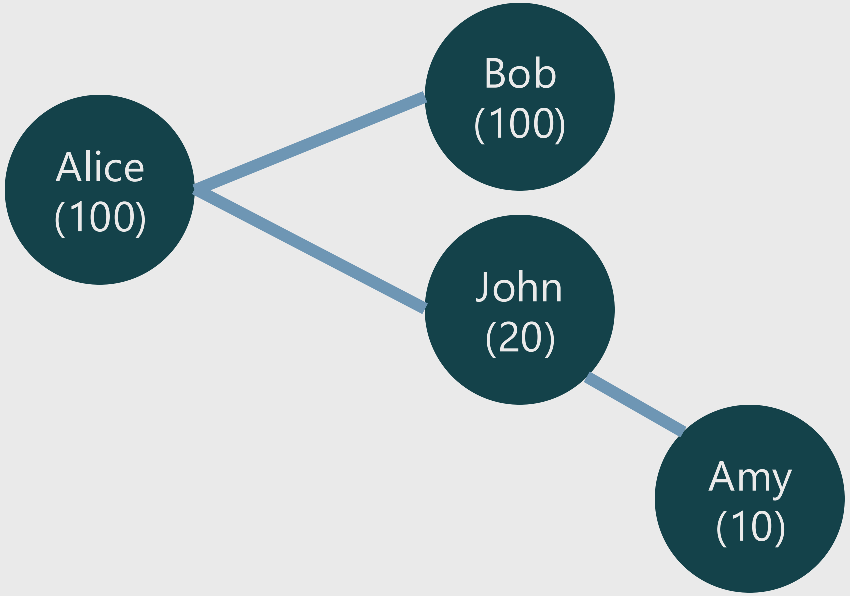
- **How to represent and describe networks**

- **What type of analyses are possible**

- **Centrality:** Who are the key actors in the network?

Basic definitions

Networks (graphs)



Nodes (vertices, actors) connected by **edges** (links, connections, relationships)

N: **Nodes** = {Alice, Bob, John, Amy}

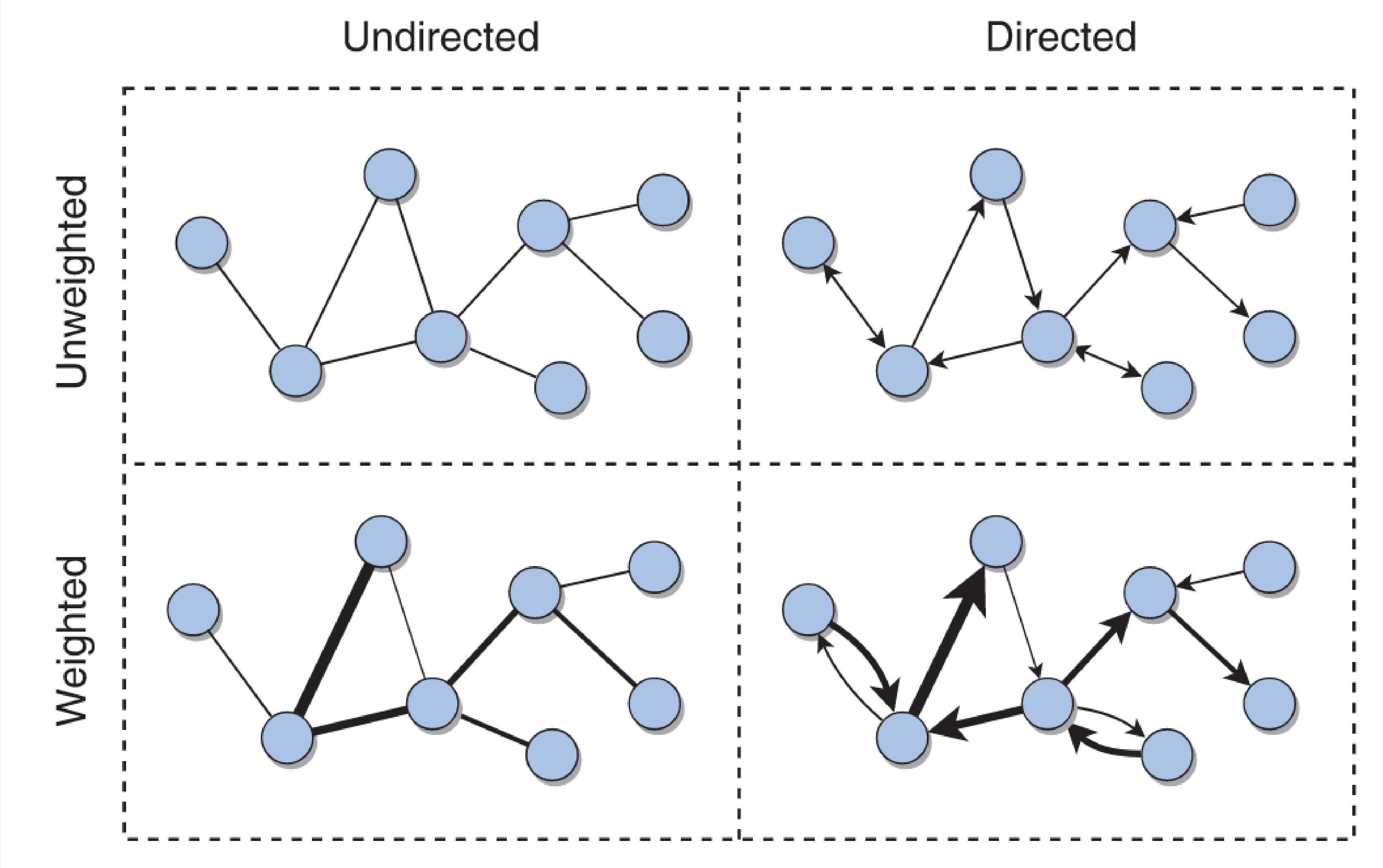
E: **Edges** = {(Alice, Bob), (Alice, John), (John, Amy)}

The edge (i,j) connects node i to node j

Nodes can have **attributes** (e.g. gender, income, etc)

Edges can have **attributes** (e.g. type, strength, etc)

Directed vs undirected; weighted vs unweighted



Source: A first course in network science (2020)

Undirected: The link (i,j) connects node i to node j in both directions

Directed: The link (i,j) connects node i (source) to node j (target)

Weighted: There is a weight associated to each edge

Degree in undirected networks

Definition: Number of neighbors in the network

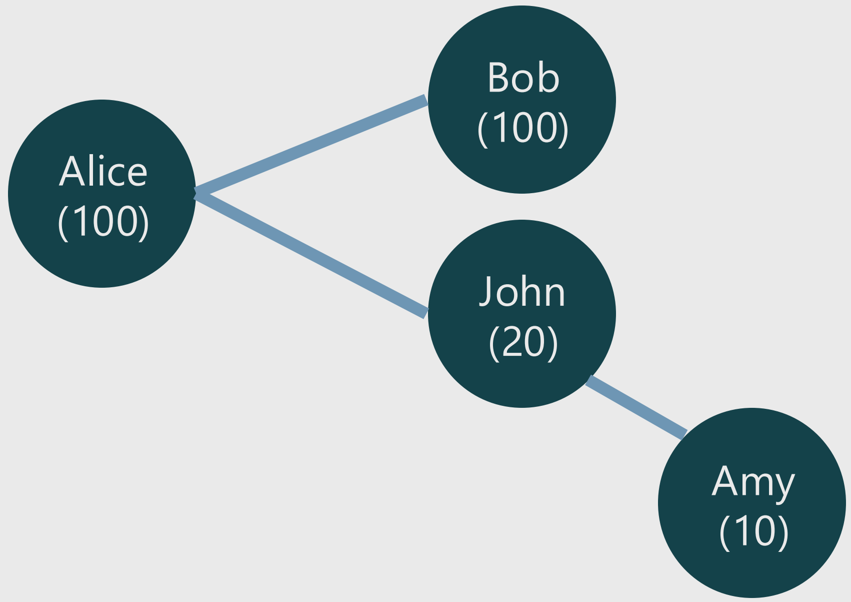
Node: degree

Alice: 2

Bob: 1

John: 2

Amy: 1



Degree in directed networks

Out-degree: Number of outgoing edges

In-degree: Number of incoming edges

Total degree: Sum of out and in degree

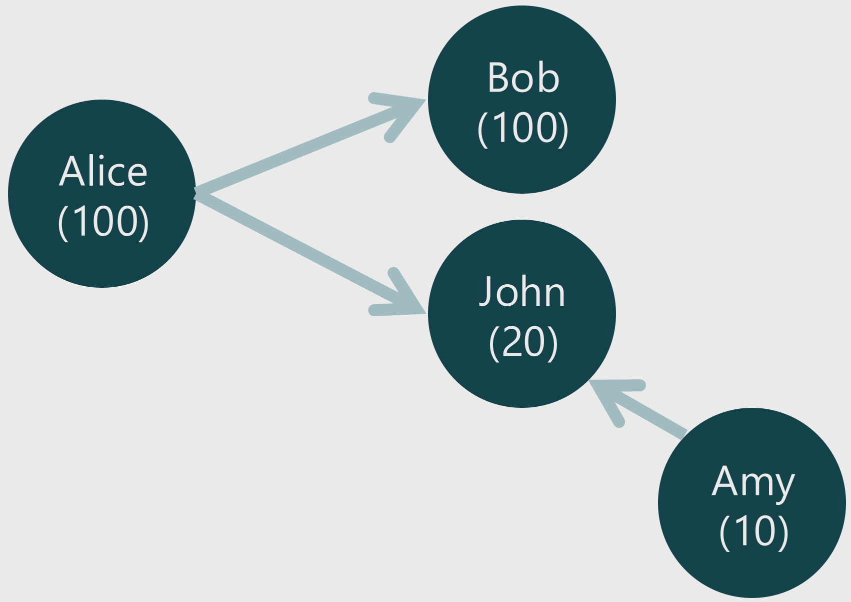
Node: (out, in, total)

Alice: (2, 0, 2)

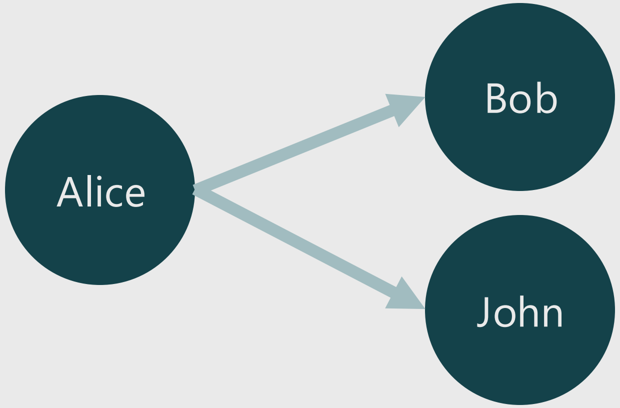
Bob: (0,1, 1)

John: (0,2, 2)

Amy: (1,0, 1)



Network representation



Adjacency list (edgelist):

- Adv: It is dense: Only keeping edges
- Disadvantage: Hard to work with

Origin	Target	Weigth
Alice	Bob	1
Alice	John	1

Adjacency matrix:

- Adv: Linear algebra is easy
- Disadvantage: It is sparse (mostly zeros).1E6 nodes → 1 trillion options

Target → ↓ Source	Alice	Bob	John
Alice	0	1	1
Bob	0	0	0
John	0	0	0

In computer → Sparse matrices: Best of both worlds

Network formats

CSV:

- **Adjacency list:** very common
- No node features (or separate file)

GraphML:

- **XML-based**, supports:
- Support for many types of graphs/edges/edges attributes
- Not great for the web (JSON would be smaller/better)

GML (Graph Modelling Language):

- Human-readable text format
- Hierarchical structure
- No formal schema (e.g. key definition): less standard/portable

Software-specific formats (e.g. Pajek)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<graphml xmlns="http://graphml.graphdrawing.org/xmlns"
         xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
         xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns http://graphml.
         graphml.xsd">

  <!-- property keys -->
  <key id="key0" for="node" attr.name="_pos" attr.type="vector_float" />
  <key id="key1" for="graph" attr.name="citation" attr.type="string" />
  <key id="key2" for="graph" attr.name="description" attr.type="string" />
  <key id="key3" for="edge" attr.name="layer" attr.type="short" />
  <key id="key4" for="graph" attr.name="layer_key" attr.type="python_object" />
  <key id="key5" for="node" attr.name="name" attr.type="string" />
  <key id="key6" for="graph" attr.name="name" attr.type="string" />
  <key id="key7" for="node" attr.name="nodeLabel" attr.type="string" />
  <key id="key8" for="graph" attr.name="tags" attr.type="vector_string" />
  <key id="key9" for="graph" attr.name="url" attr.type="string" />
  <key id="key10" for="edge" attr.name="weight" attr.type="boolean" />
```

```xml
  <!-- vertices -->
  <node id="n0">
    <data key="key0">3.9477401370487888, 1.2719081581250928</data>
    <data key="key5">ACCIAIUOL</data>
    <data key="key7">ACCIAIUOL</data>
  </node>
  <node id="n1">
    <data key="key0">4.0067822619703701, 1.2437653936176905</data>
    <data key="key5">ALBIZZI</data>
    <data key="key7">ALBIZZI</data>
  </node>
```

```xml
  <!-- edges -->
  <edge id="e0" source="n0" target="n8">
    <data key="key3">1</data>
    <data key="key10">1</data>
  </edge>
  <edge id="e1" source="n1" target="n5">
    <data key="key3">1</data>
    <data key="key10">1</data>
  </edge>
```

Network visualization

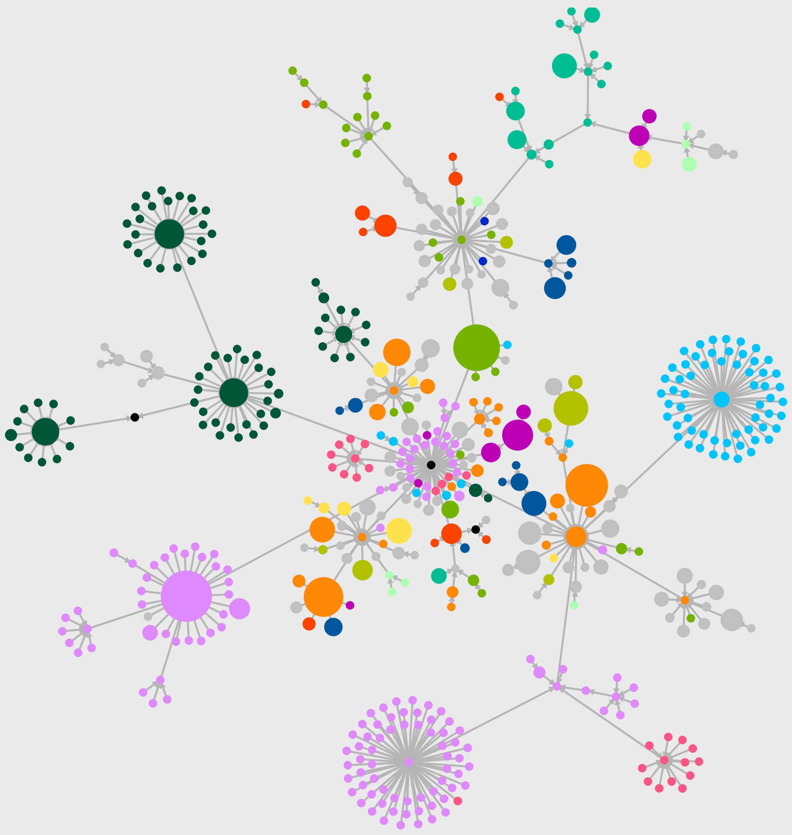
Layouts: Determine the positions of the nodes in the plot

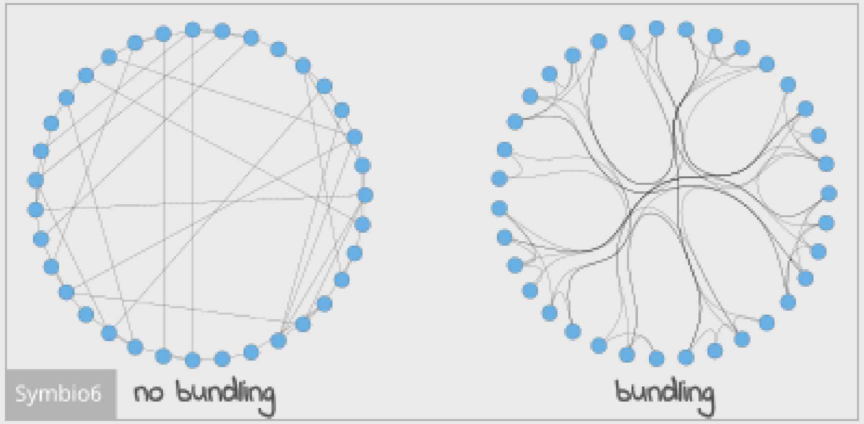
Spring layout (force-directed/Fruchterman-Reingold):

- Most used
- Connected nodes attract each other, non-connected nodes repel each other

Many more layouts, such as:

- Parametric layouts: Based on node attributes (e.g. the geometric coordinates)
- Circular layout: In a circle, avoiding edge crossing





Practical 1, exercise 1

If you don't have experience with Python:

- Click on Colab: https://net-science.github.io/

If you have experience with Python:

1. Download materials: github.com/jgarciab/NetworkScience (click on code -> Download Zip)
2. Extract ZIP
3. Set up Python. On **Windows & Mac** using the graphical interface:

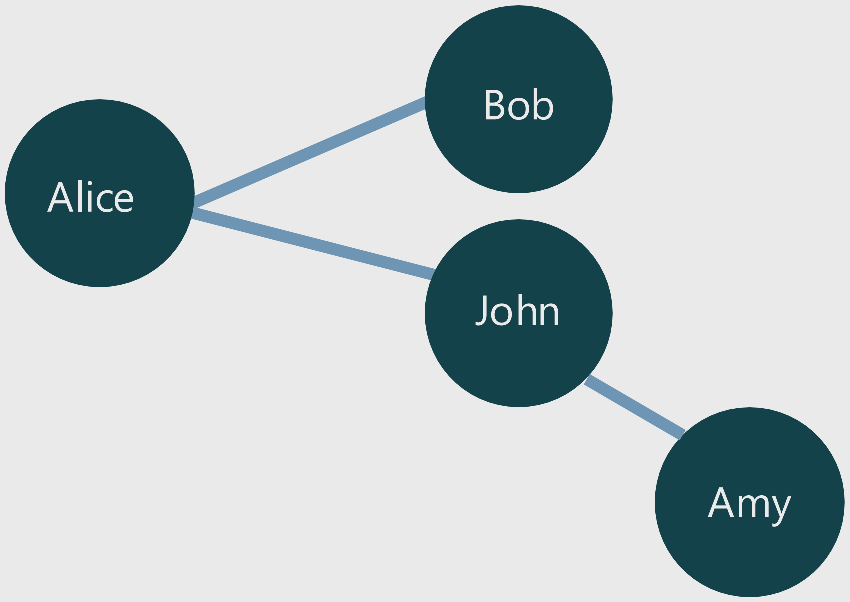
- Open Anaconda
- Go to "Environments" (left menu)
- Click on "Import" and specify the file "environment.yml" (it's one of the files that you downloaded)
- Activate environment by clicking in the "play" button next to the environment.

On **Linux** (or Mac)

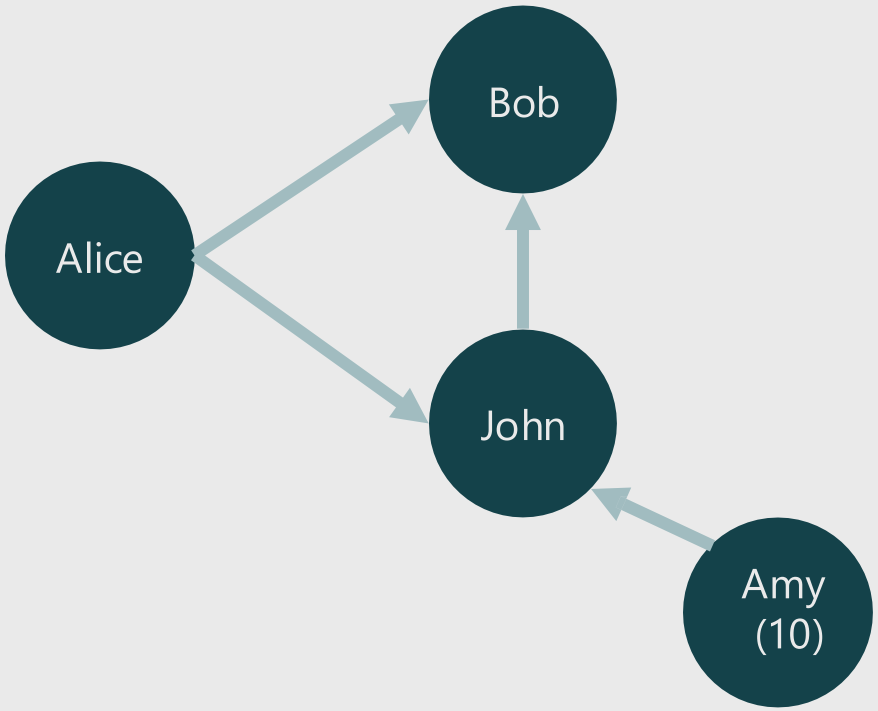
- Open a conda terminal (open a terminal)
- Navigate to the directory with the code using dir (ls) to list the files and cd XXX (cd XXX) to enter directory XXX.
- Create a new environment: *conda env create -f environment.yml*
- Activate environment: *conda activate networks*
- Launch jupyter notebook: *jupyter notebook*

4. Do exercise 1 in: day*1a_intro_networks.ipynb*

Types of networks



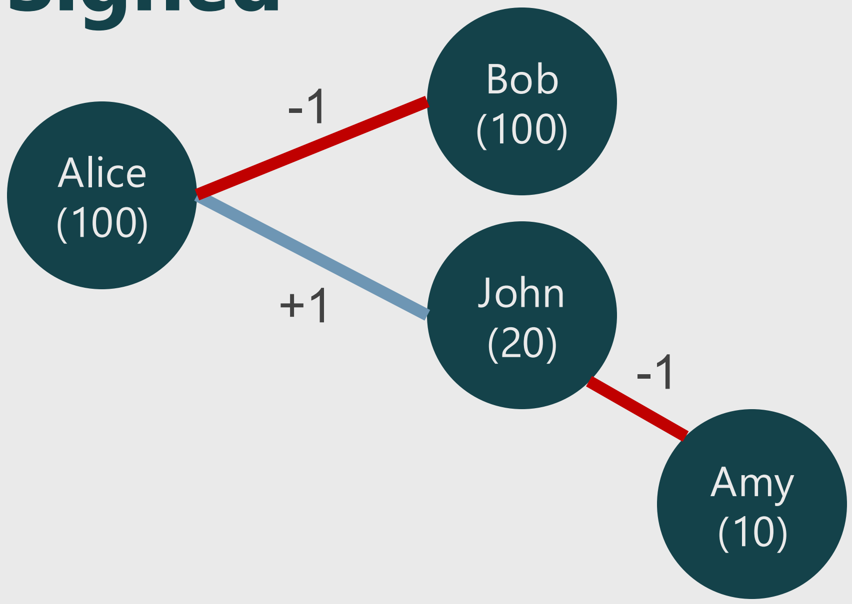
Undirected Acyclic Graphs (Trees)

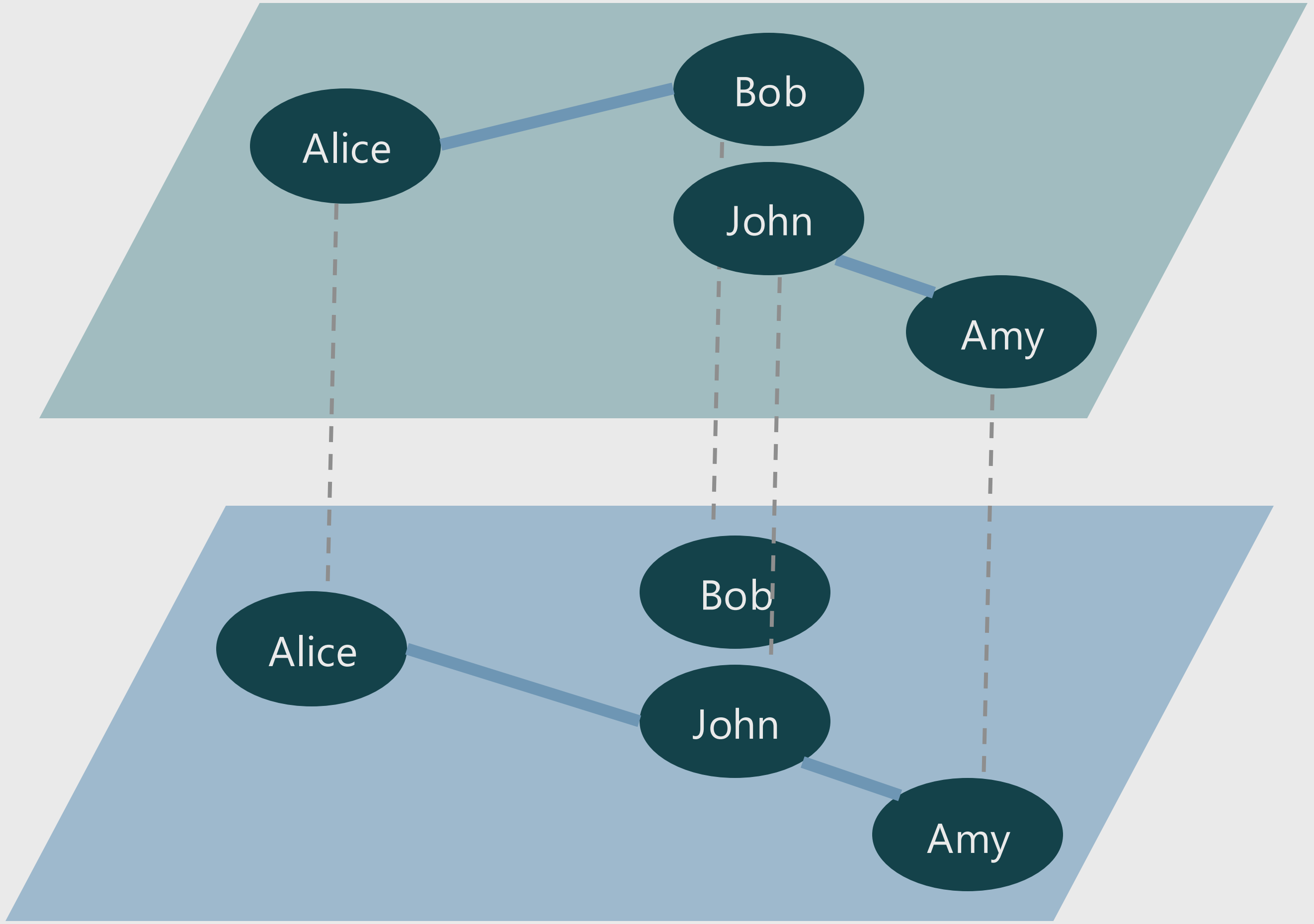


Directed Acyclic Graphs (DAGs)

Other types of networks: Multiplex

Signed

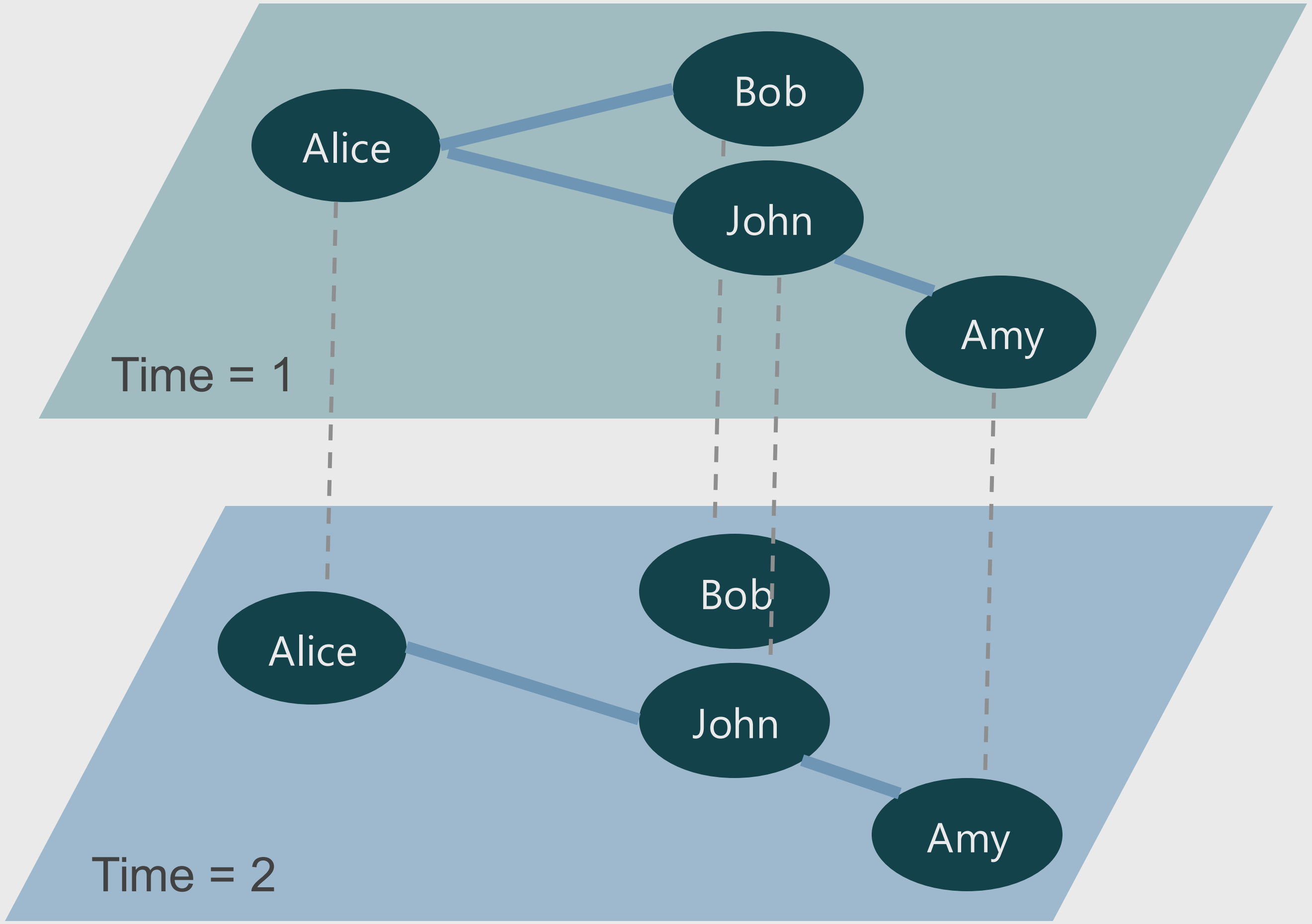




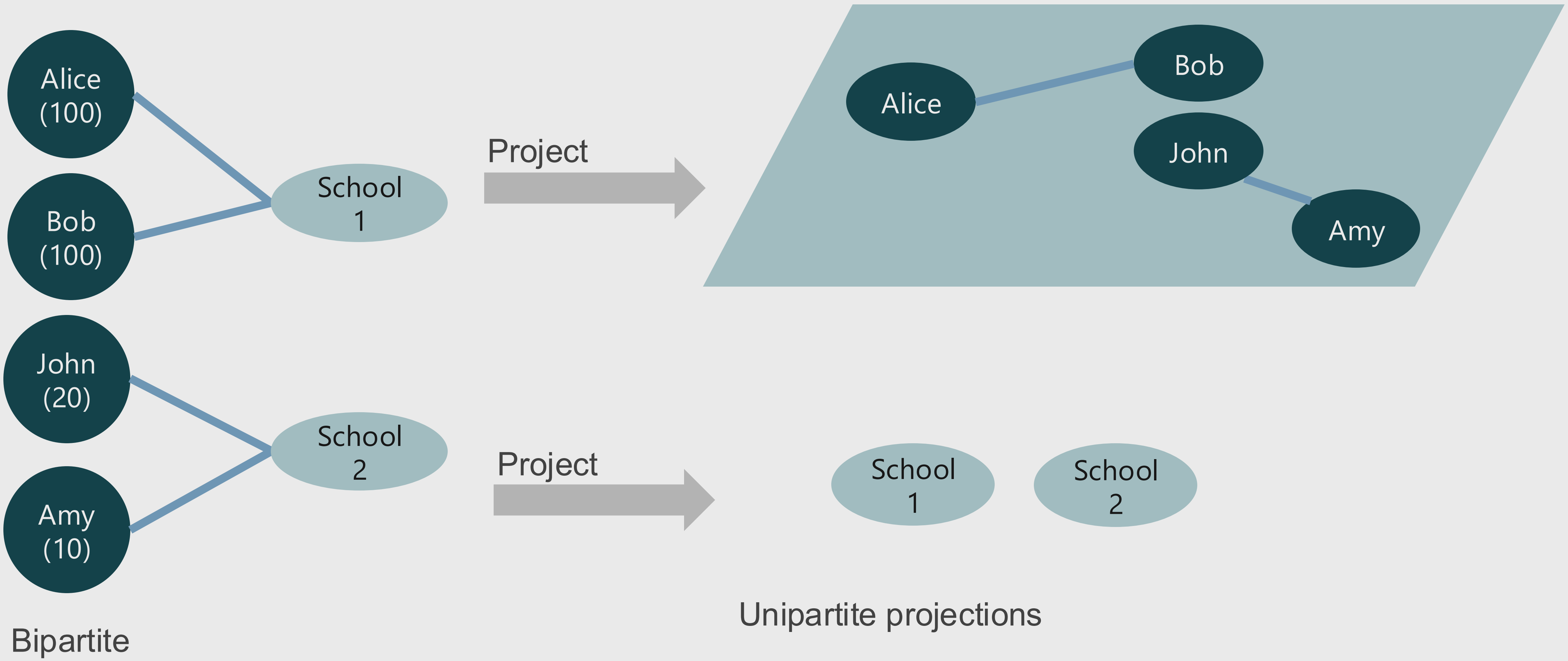
Other types of networks: Temporal

Either:

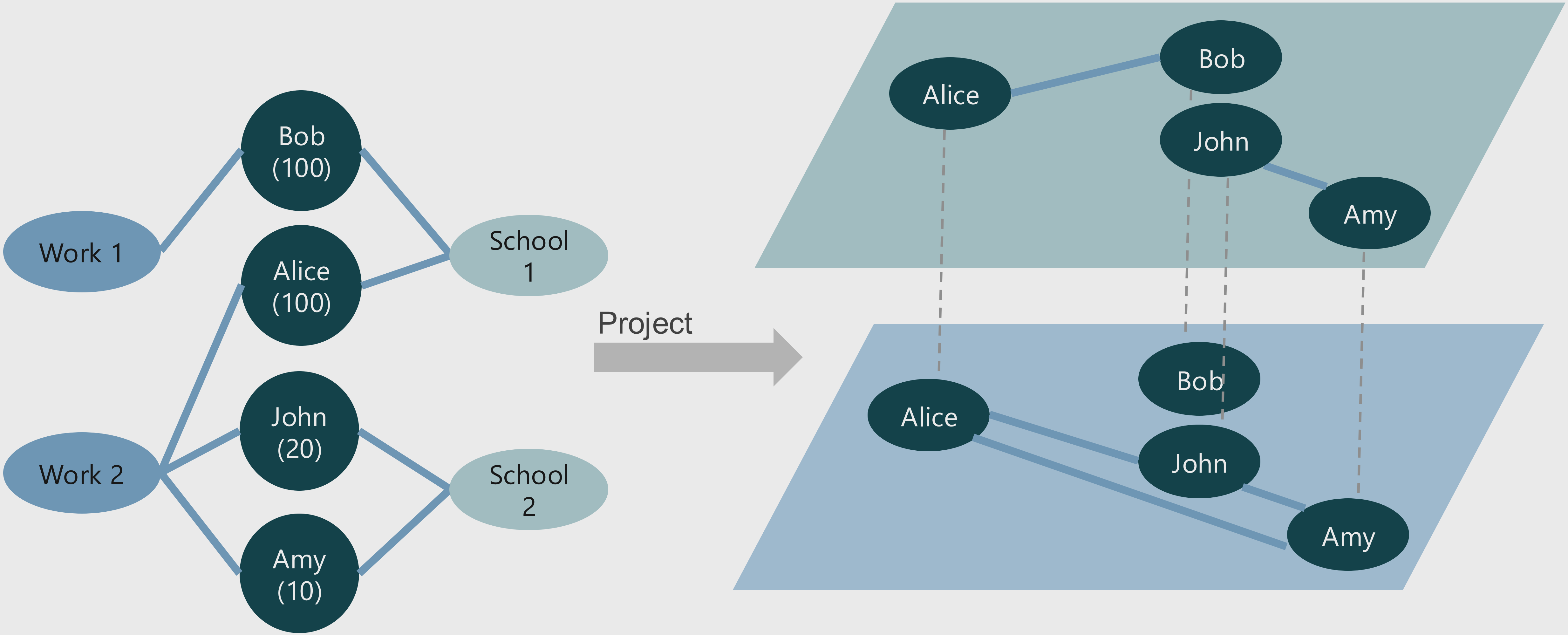
- Snapshots
- Time of events



Other types of networks: Bipartite



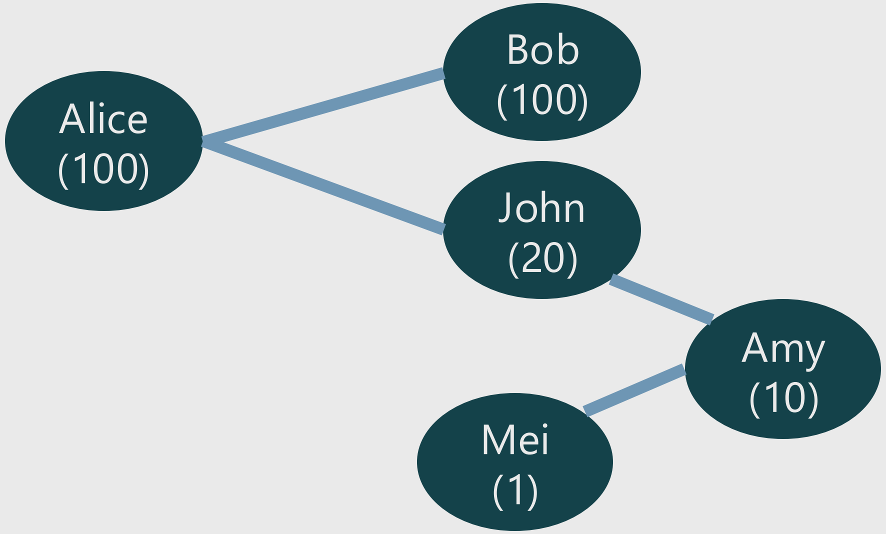
Other types of networks: Multipartite

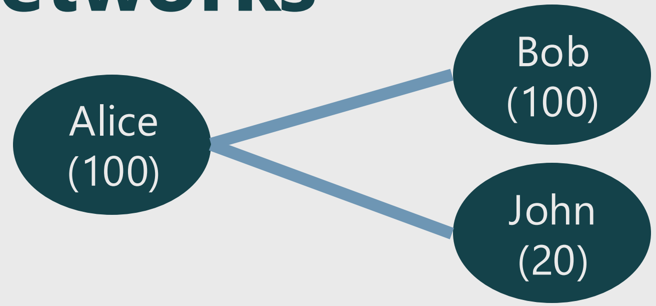


Multipartite network

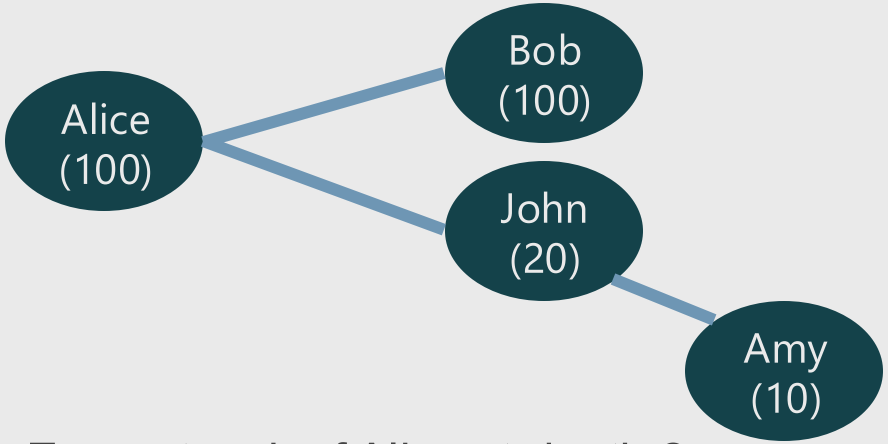
Multiplex projection

Other types of networks: Ego-networks



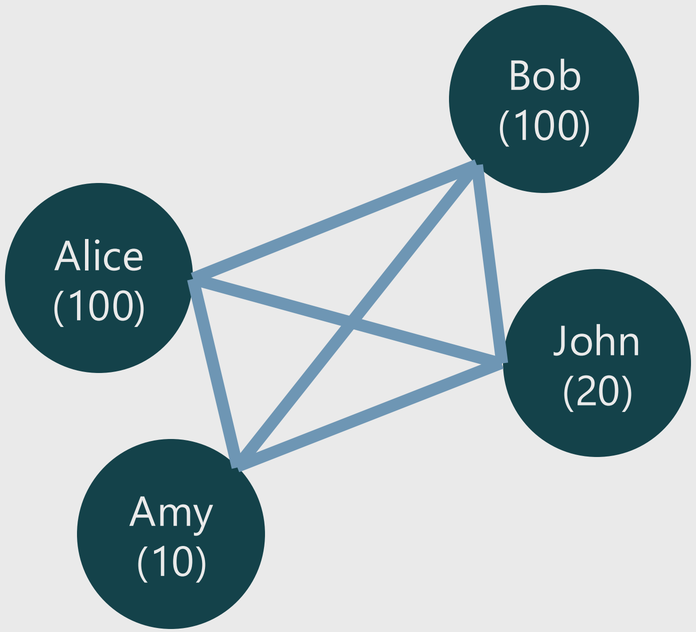


Ego network of Alice at depth 1



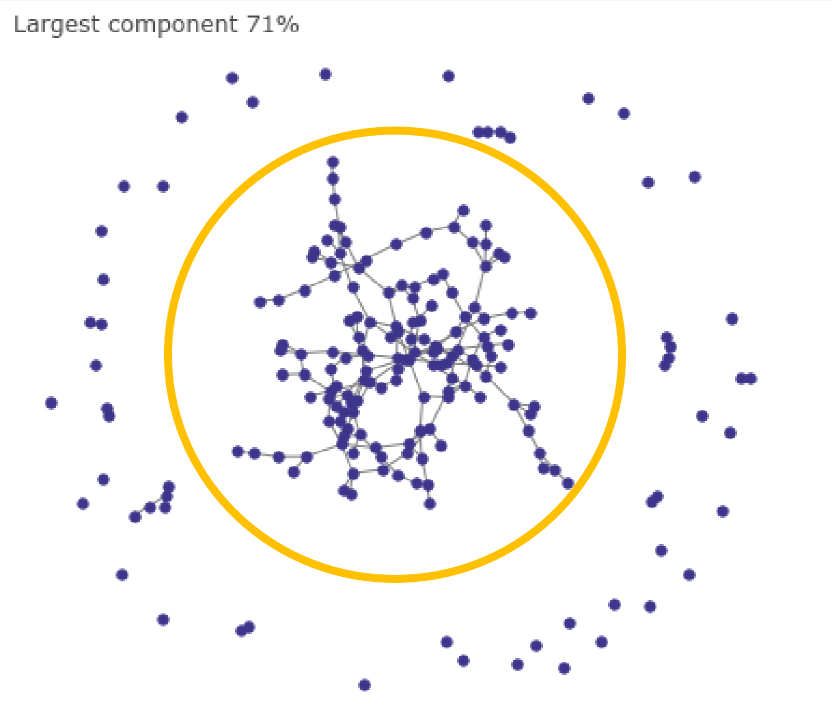
Ego network of Alice at depth 2

Other types of networks: Clique



Network characteristics

Connectedness



Real networks are typically connected, forming a **“giant component”**

If the average degree < 1 → many small components

If the average degree > 1 → suddenly the system becomes connected

Let’s try this!

Small world: six degrees of separation

Milgram's experiment (1967)

Image source: Drewonwiki, Wikipedia

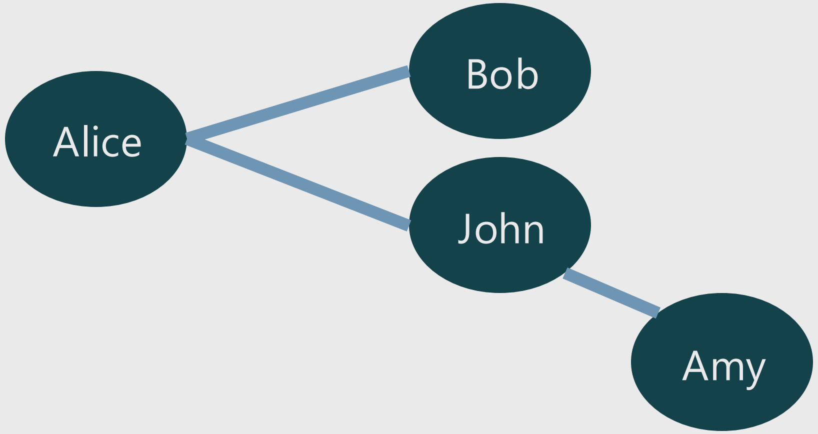
Shortest path between node 1 and node 2:

- Minimum number of steps requires to go from node 1 to node 2
- Between Alice, Amy → 2

Diameter:

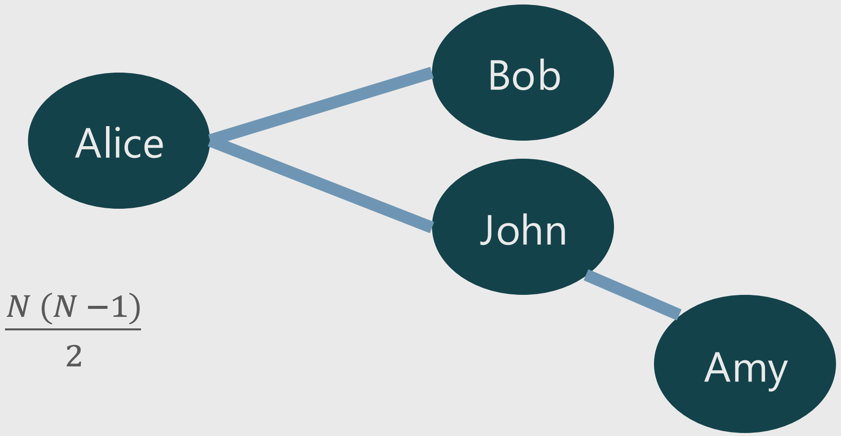
- Longest "shortest path" between two nodes
- In our network: 3 (Bob -> Alice -> John -> Amy)

Real networks have **small diameters** because hubs connect diverse parts of the network



Density

Definition: Number of edges present / potential number of edges



- Number of edges = 3
- Potential number of edges in directed network = (4*3)
- Potential number of edges in undirected network = (4*3)/2 $= \binom{N}{2} = \frac{N\,(N-1)}{2}$

Density = 3/6 = 50%

Real networks are typically **sparse**

As size increases density decreases (average degree is usually fixed)

Practical 1, exercise 2.1–2.3 (included)

Compare the PPI network with the Twitter network (ic2s2_netsci_3.tsv).

What characteristics apply to both? Which don’t? Why?

Connectedness:
- Twitter: 1 component, 1554 nodes
- PPI: 108 components, 832 nodes in largest

Diameter:
- Twitter: 6, 2.5 average path
- PPI: 17, 5.7 average path

Density,
- Twitter: 2.5%,
- PPI: 0.3%,

Local and global clustering

Clustering // Holland & Leinhardt, (1971); Strogatz & Watts (1998):

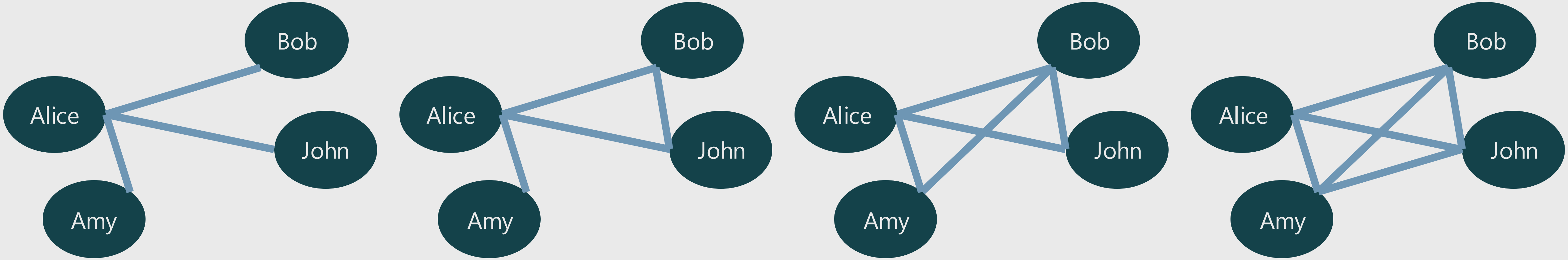
- The share of your connections who are themselves connected to each other
- The share of triads (connected sets of 3 nodes) that are closed (form a triangle)

Local clustering = For each node. Unit of observation = node

Global clustering (transitivity) = For the network. Unit of observation = triad

Real networks have **high clustering**

Clustering of Alice:

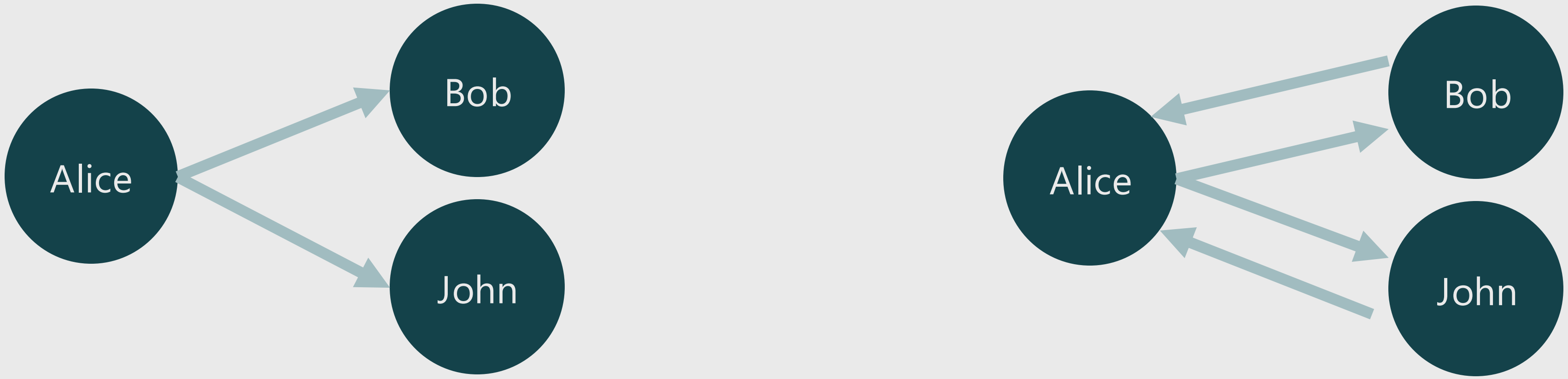


0/3 1/3 2/3 3/3

Reciprocity

Only for directed networks

Ratio of the number of edges pointing in both directions to the total number of edges in the graph.



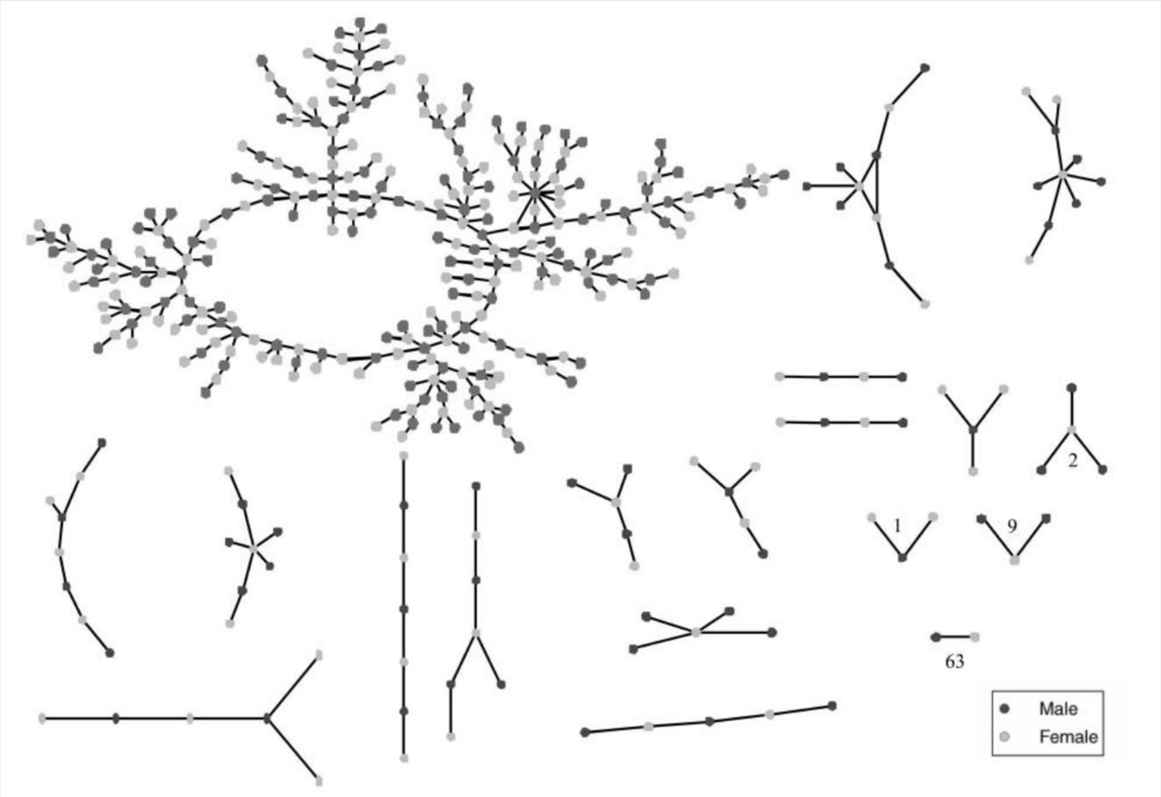
Assortativity (homophily)

Preference for nodes to attach to others that are similar in some way

Defined with respect of an attribute (e.g. gender)

Ranges from -1 (fully disassortative) to 1 (fully assortative)

Paraisópolis favela and Morumbi, in São Paulo
Photography by Tuca Vieira (the guardian)



Romantic links between teenagers
Bearman, Moody, Stovel (1991)

Assortativity (homophily)

At the network level:

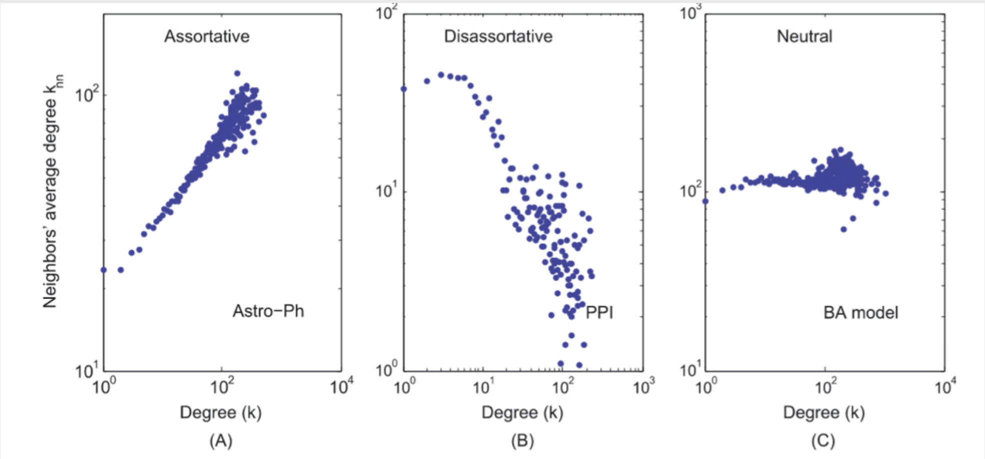
- Continuous variable: Pearson's correlation across edges.
- Categorial unordered variable = Modularity
  - (Actual links between edges between nodes of same type – expected number of links between nodes of same type)/number of links

Mixing patterns in networks, Newman, Physical Review E, 67 026126, 2003

At the local level:

- Real networks can be locally assortative or disassortative
- Exercise: Draw a degree-assortative network

Multiscale mixing patterns in networks, Peel, Delvenne and Lambiotte (2018)



Jiang et al (2016)

Practical 1, exercise 2.4–2.6 (included)

Compare the PPI network with the Twitter network (ic2s2_netsci_3.tsv).

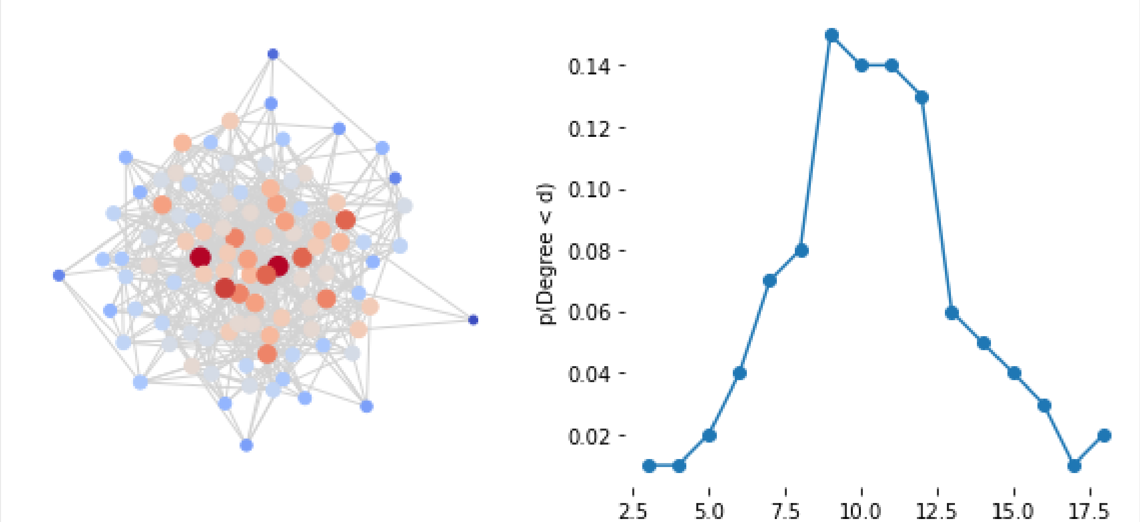
What characteristics apply to both? Which don't? Why?

Transitivity:
- Twitter: 21%,
- PPI: 1.5%,

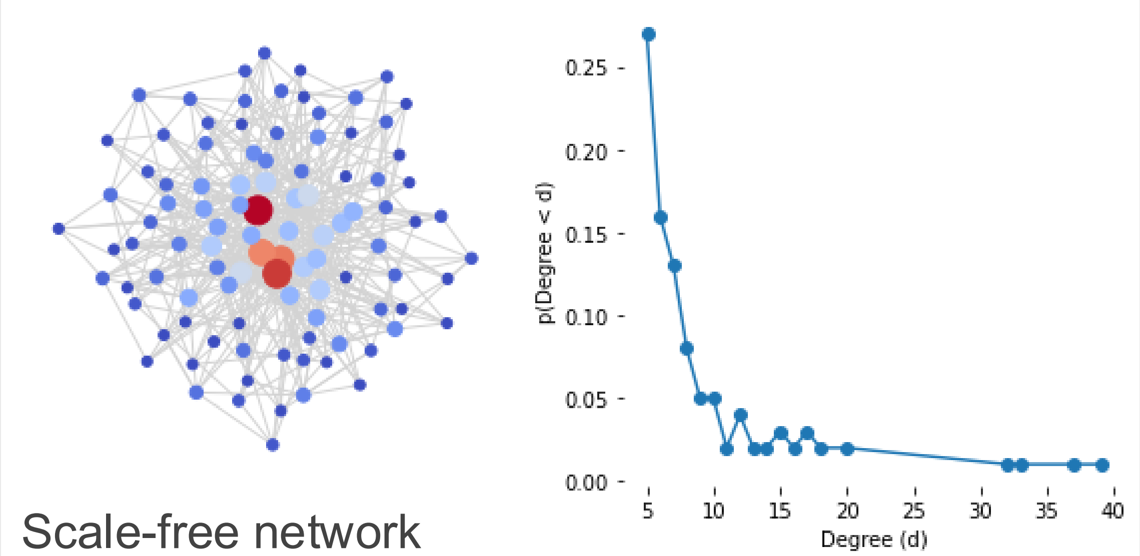
Average clustering:
- Twitter: 35%,
- PPI: 4.3%,

Degree assortativity:
- Twitter: -3.2%,
- PPI: -16.7%,

Heavy tails / scale-free



Random network



Scale-free network

Networks are not random, they have heavy tails

PDF (probability density function)

→ Degree vs probability of degree

→ Represented by histogram

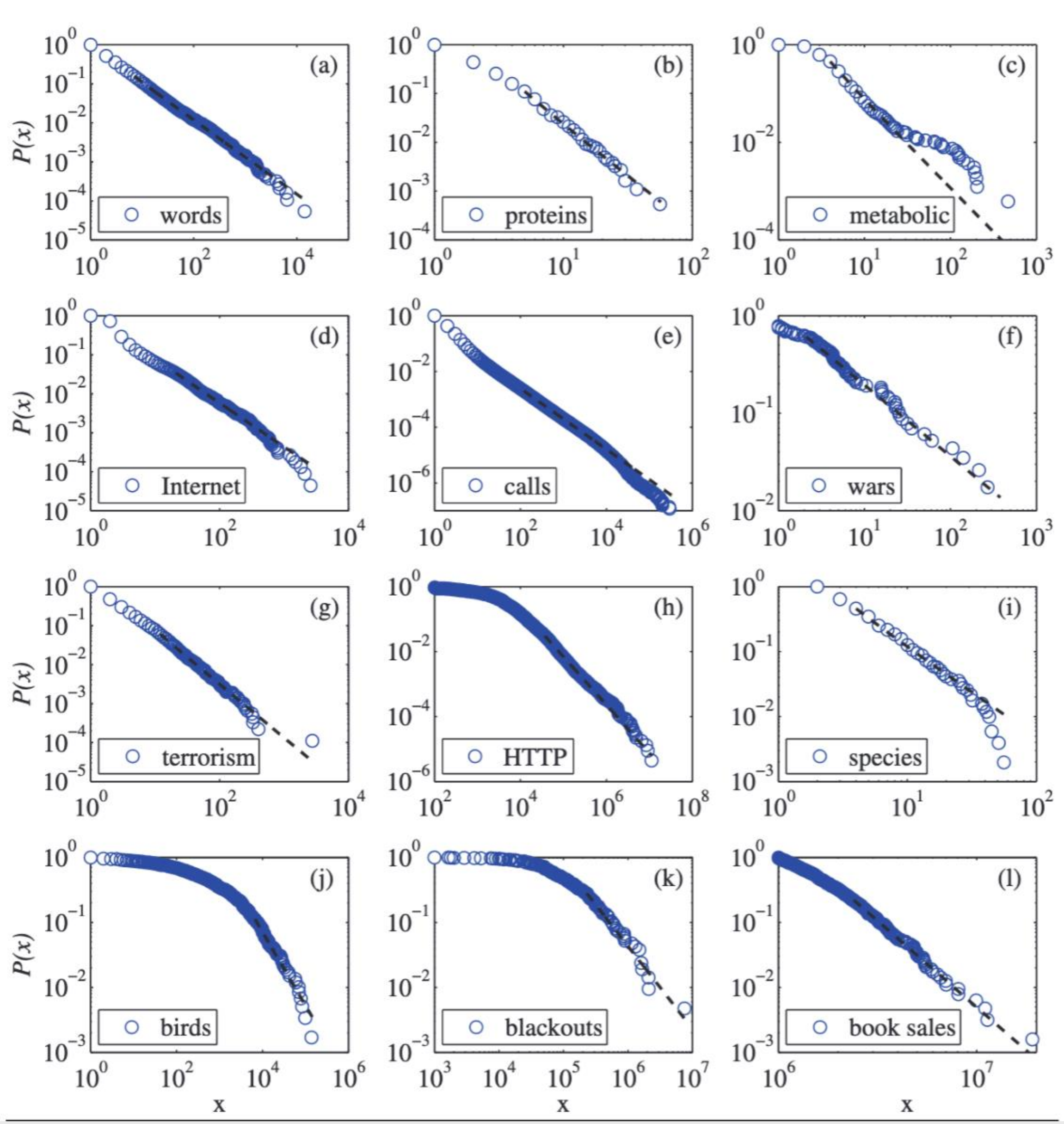
Many possible mechanisms:

- Multiplicative growth
- Preferential attachment (Rich get richer, Matthew effect)
- Copying models

Growing networks: https://www.stat.cmu.edu/~cshalizi/networks/16-1/lectures/08/li.pdf

Heavy tails

Most complex systems have **heavy tail distributions**



Clauset, Shalizi & Newman (2009)

Random networks don't have heavy tails

PDF (probability density function)

→ Degree vs probability of degree

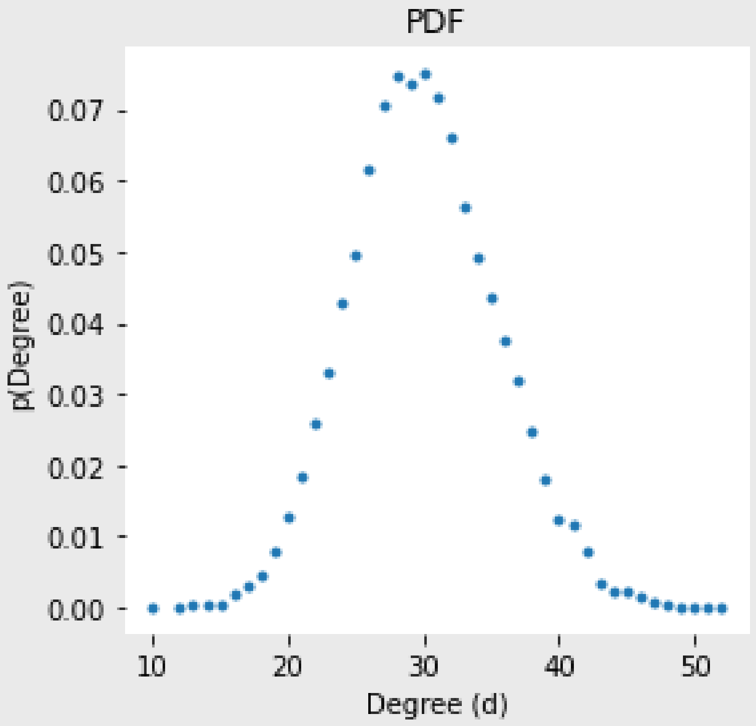
→ Represented by histogram

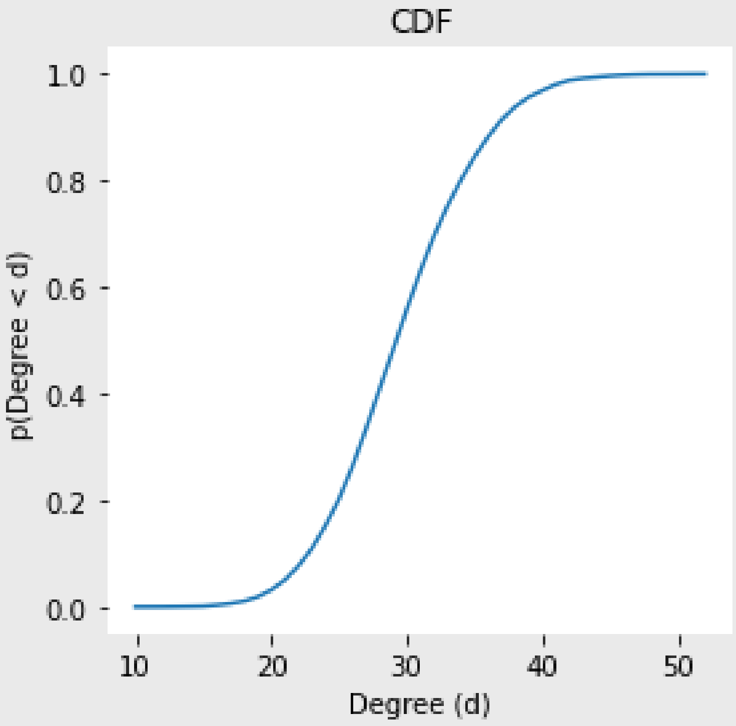
CDF (cumulative density function)

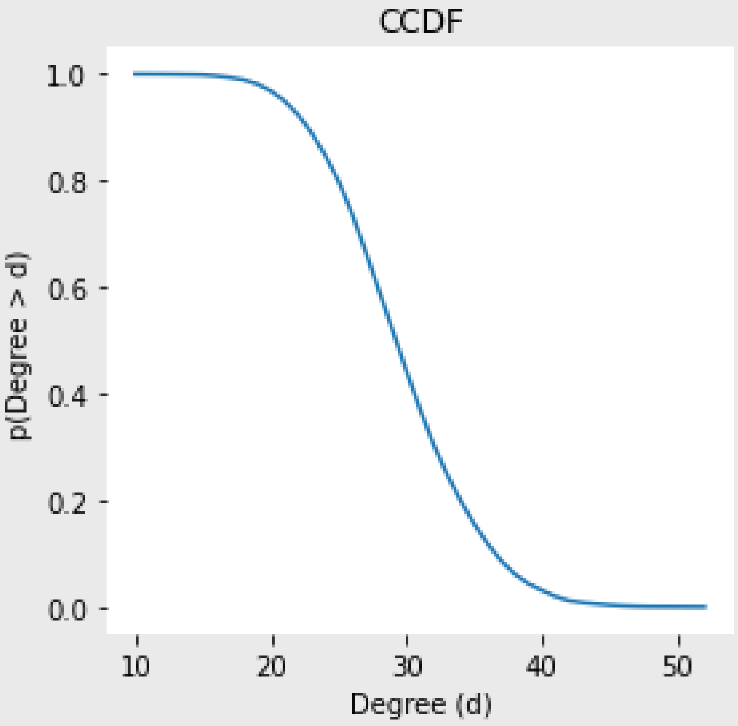
→ Degree s vs probability degree < s

CCDF: Complementary CDF

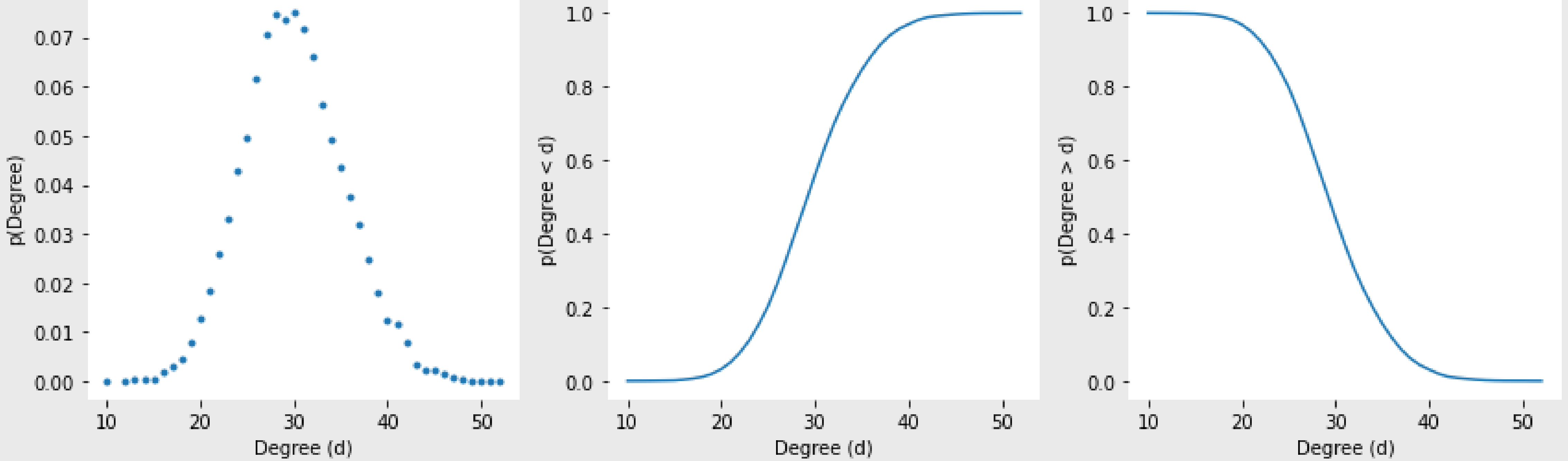
→ Degree s vs probability degree > s



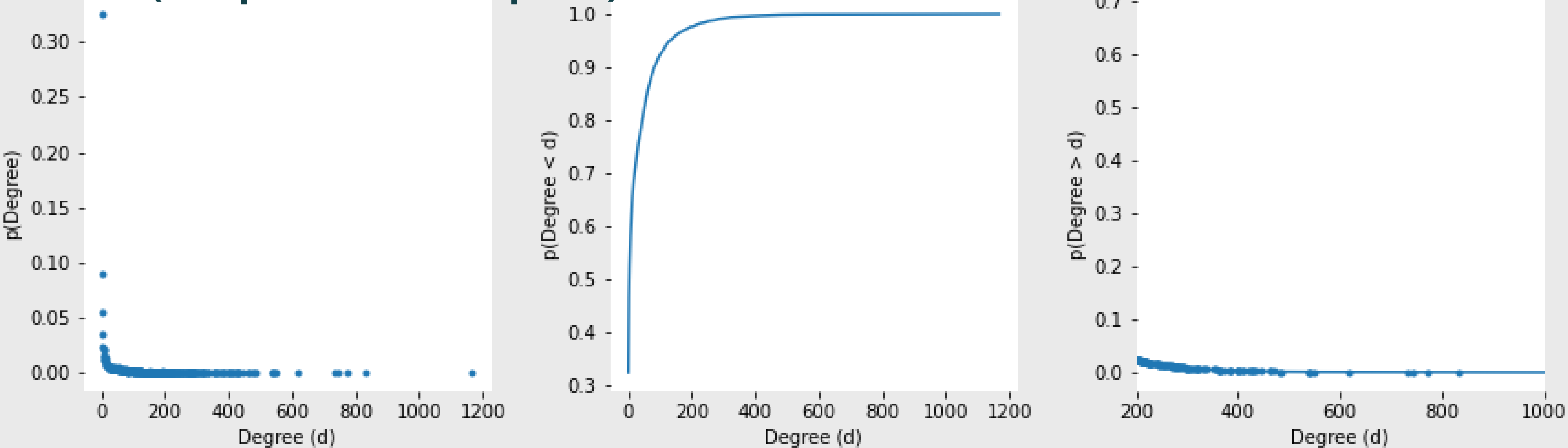




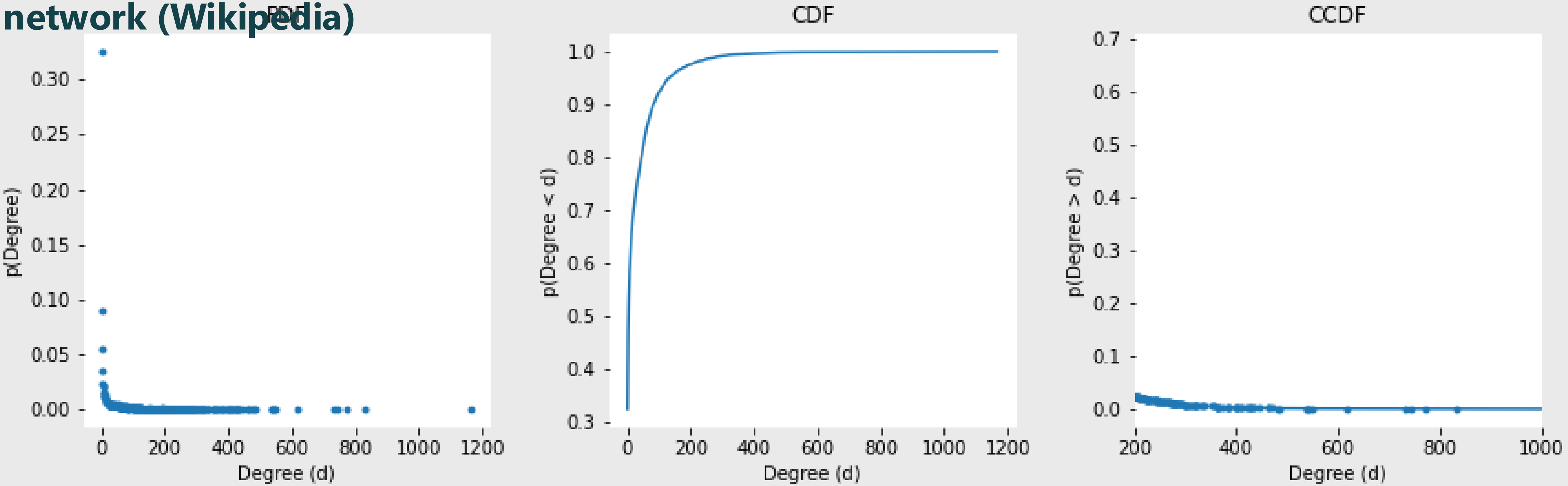
Random network



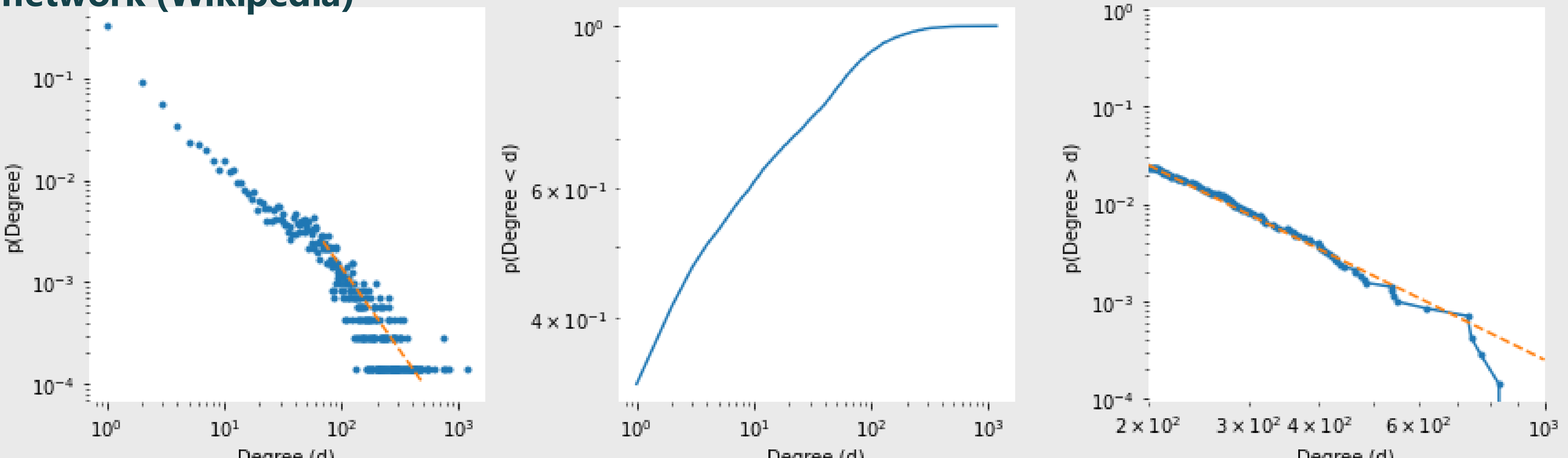
Real network (Wikipedia admin requests)



Real network (Wikipedia)



Real network (Wikipedia)



Is it a power-law? $P(d) \sim d^{-\alpha}$

Critical Truths About Power Laws

Most reported power laws lack statistical support and mechanistic backing.

MICHAEL P. H. STUMPF AND MASON A. PORTER

SCIENCE • 10 Feb 2012 • Vol 335, Issue 6069 • pp. 665-666 • DOI: 10.1126/science.1216142

Scale-Free Networks

Scientists have recently discovered that various complex systems have an underlying architecture governed by shared organizing principles. This insight has important implications for a host of applications, from drug development to Internet security

BY ALBERT-LÁSZLÓ BARABÁSI AND ERIC BONABEAU

50 SCIENTIFIC AMERICAN MAY 2003

Article | Open Access | Published: 04 March 2019

Scale-free networks are rare

Anna D. Broido & Aaron Clauset

Nature Communications **10**, Article number: 1017 (2019)

Love is All You Need

Clauset's fruitless search for scale-free networks

by Albert-László Barabási, March 6, 2018

Comment | Open Access | Published: 04 March 2019

Rare and everywhere: Perspectives on scale-free networks

Petter Holme

Nature Communications **10**, Article number: 1016 (2019) | Cite this article

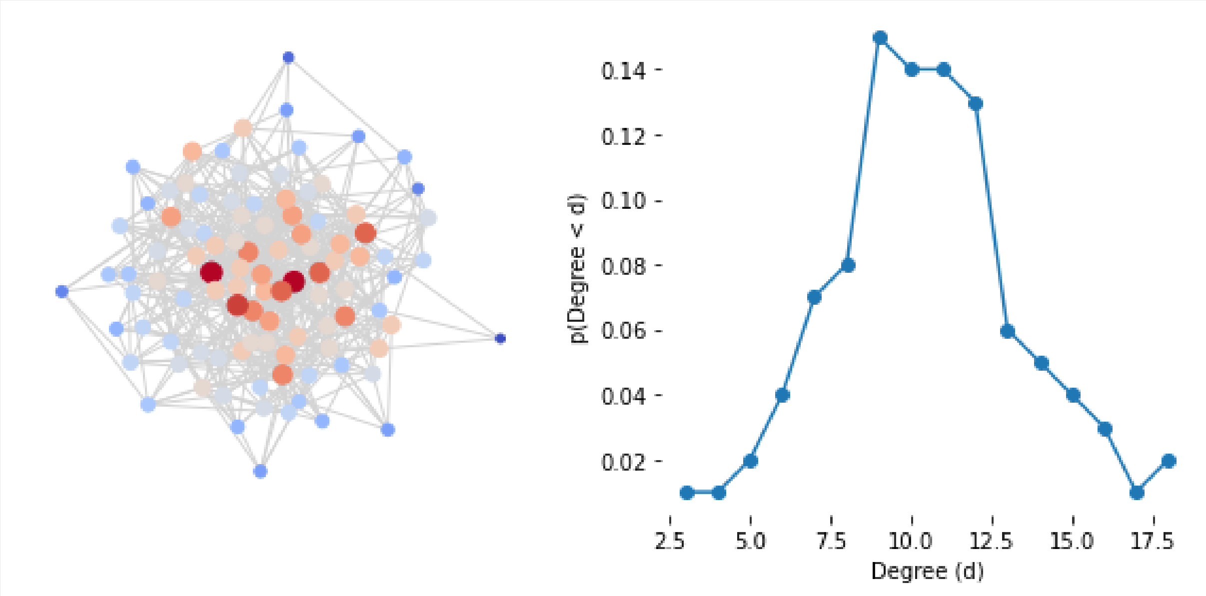
True scale-free networks hidden by finite size effects

Matteo Serafino, Giulio Cimini, Amos Maritan, +3 , and Guido Caldarelli Authors Info & Affiliations

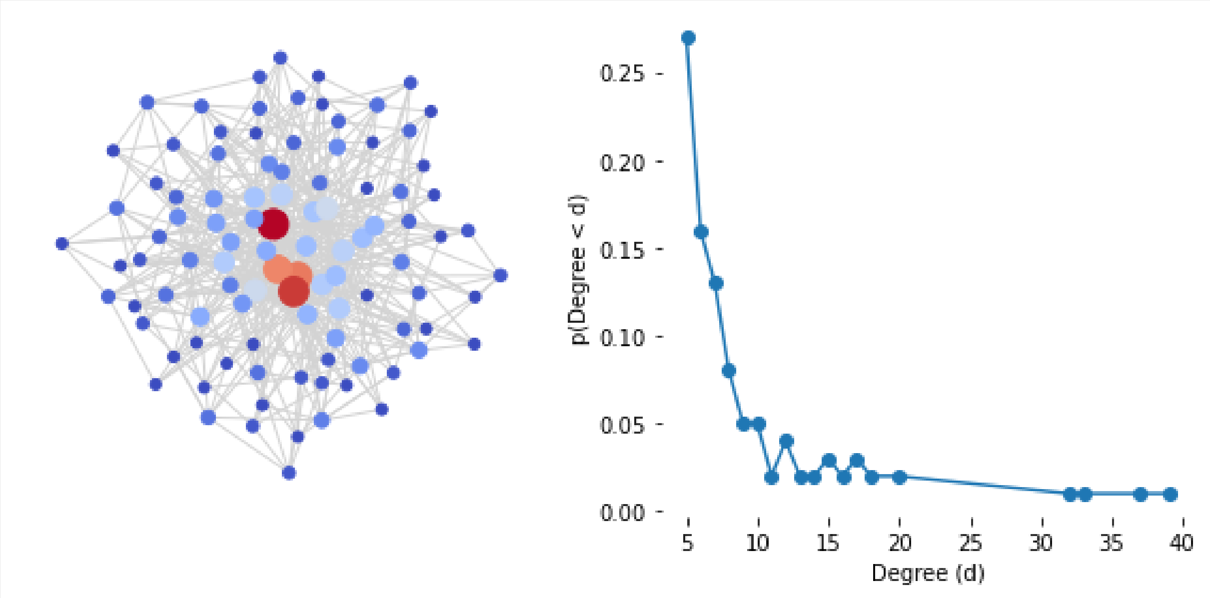
Edited by Lai-Sang Young, New York University, New York, NY, and approved November 2, 2020 (received for review July 3, 2020)

December 30, 2020 | 118 (2) e2013825118 | https://doi.org/10.1073/pnas.2013825118

Robustness to failures
Fragility to targeted attacks



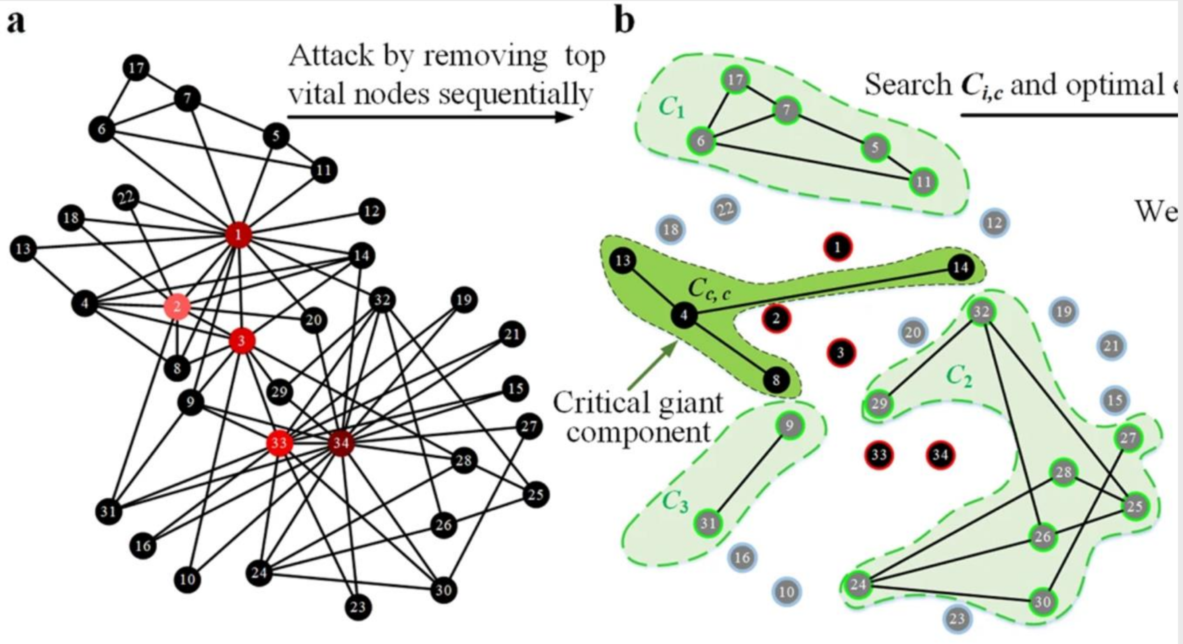
Random network



Power-law network

Robustness to failures
Fragility to targeted attacks

Albert, Jeong, Barabasi (2000) Attack and error tolerance of complex networks



Li et al (2011)

Practical 1, exercise 2.7

Compare the PPI network with the Twitter network (ic2s2_netsci_3.tsv).

What characteristics apply to both? Which don’t? Why?

Robustness:
- Twitter: quite robust, start failing at 30%
- PPI: very fragile, fails at 5%

Practical 1, exercise 3

In 3.7 compare the PPI network (default example) and the Twitter network (IC2S2).

Types of analysis

They should fit your research question

Types of analysis: Descriptive statistics

Describe the network characteristics (density, diameter, average degree, clustering, etc)

http://javier.science/panel_network/

Types of analysis: Centralities

What are the most important nodes in the network?

- The one with more connections ➔ **Degree centrality**
- The one linked to more important neighbors ➔ **Pagerank / Eigenvector / Katz centrality**
- The one closest to all other nodes ➔ **Closeness centrality**
- The ones that act as bridges between communities? ➔ **Betweeness centrality**

More on this this afternoon

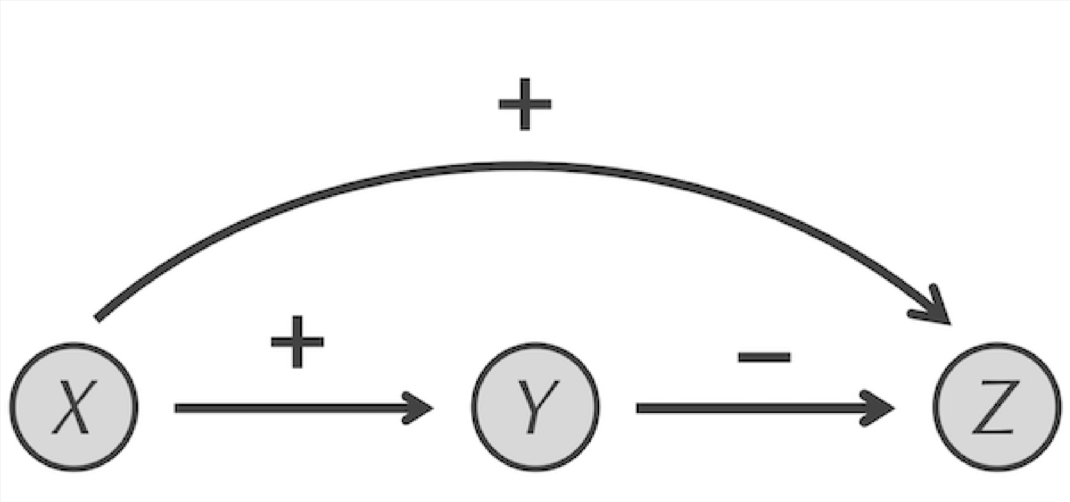
Types of analysis: Node-level regression

Calculate node-level features:

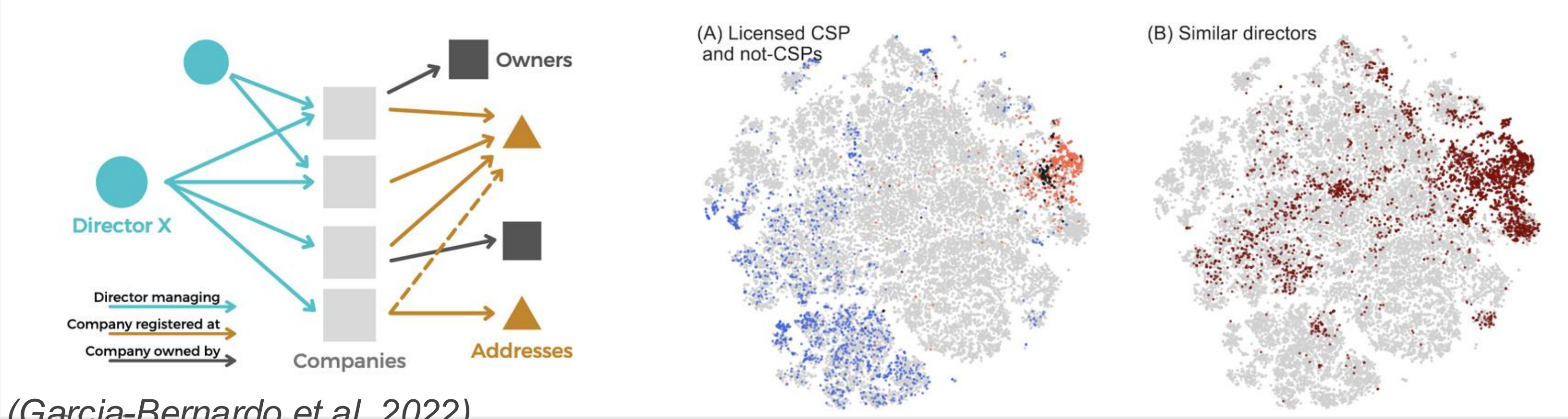
- Centrality
- Local clustering (transitivity / embeddedness)
- Local reciprocity
- Local assortativity (homophily)
- …
- Include in your prediction model (e.g. a regression)

Types of analysis: Motif detection

Find overrepresented patterns



Feed-forward loop (https://biologicalmodeling.org/motifs/feedforward)



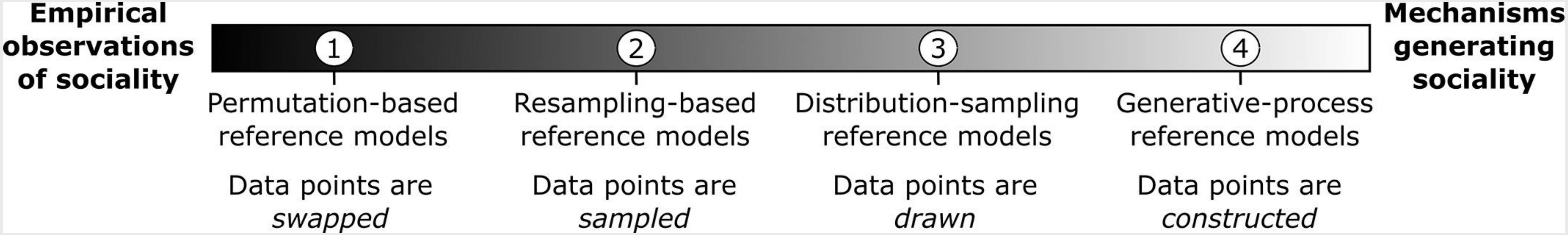
(Garcia-Bernardo et al, 2022)

Types of analysis: Testing hypothesis

We observe some behavior in the network (e.g. the clustering is 0.5). Is this relevant?

Approach: Create a reference model (see *Hobson 2021* for a great guide) to compare with it:

- *Permutation*: e.g. shuffle node attributes or permute edges (configuration model)
- *Generative*: data is constructed from rules.
    - Mechanisms (e.g. rich get richer model + homophily + triadic closure)
    - ERGM (model how the network structure depends on covariates)
    - ABM (agent-based model)



Hobson 2021

More on this on Tuesday

Permutation of attributes

Calculate significance by resampling.
Avoids running a regression (which is difficult with all the interdependences)

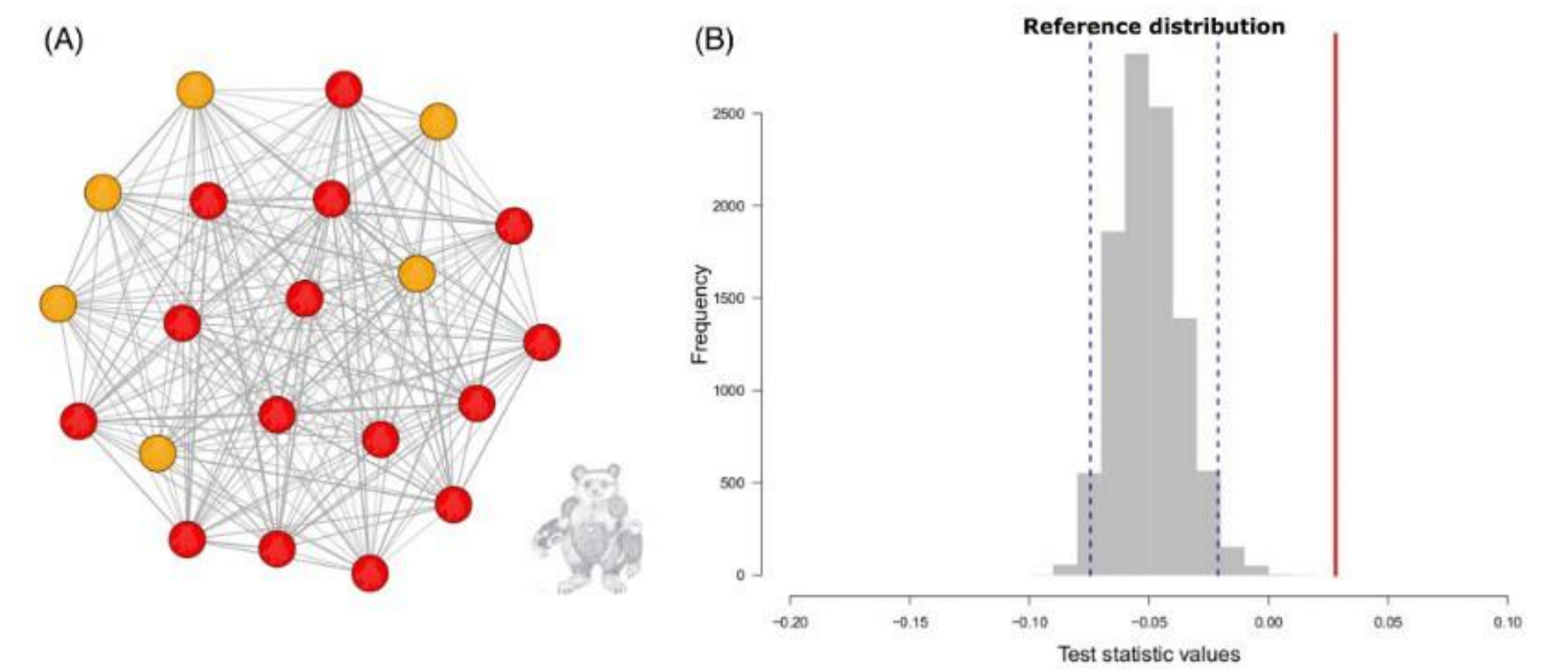
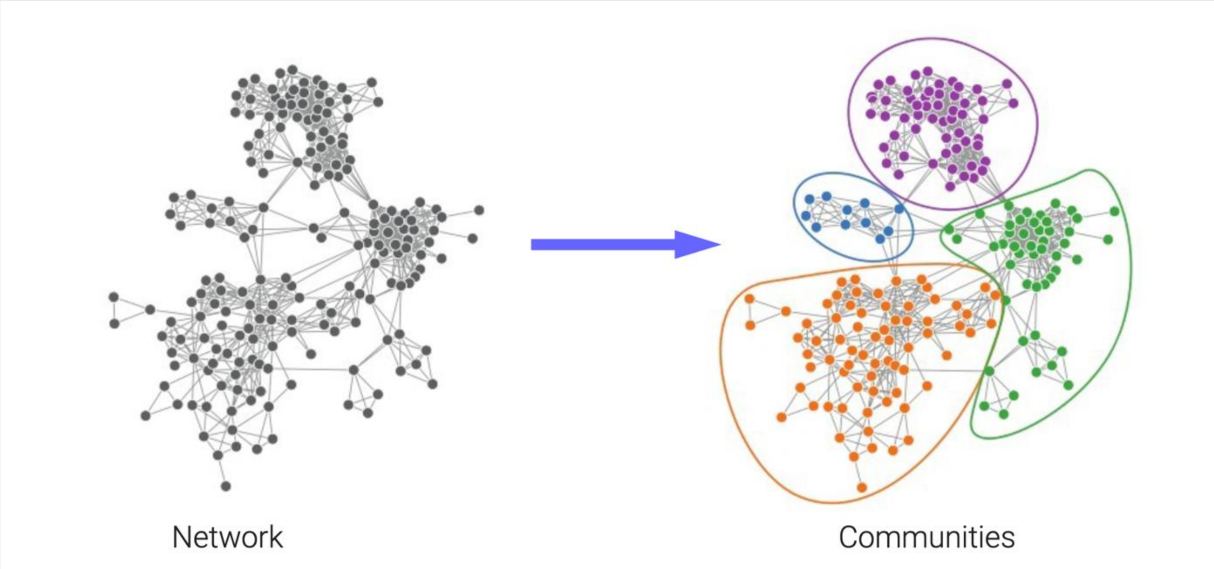


Fig 2. An example of study approaches: do burbils socially assort by nose colour? (A) Association network of burbils, with nodes colour-coded by nose colour and (B) distribution of values based on the permutation procedure of Team 1; observed value of the test statistic shown as a red solid line and the 2.5 and 97.5% quantiles of the reference distribution as blue dashed lines.

Hobson 2021

Types of analysis: Community detection



Peel et al

What clusters of nodes can we find in the network?

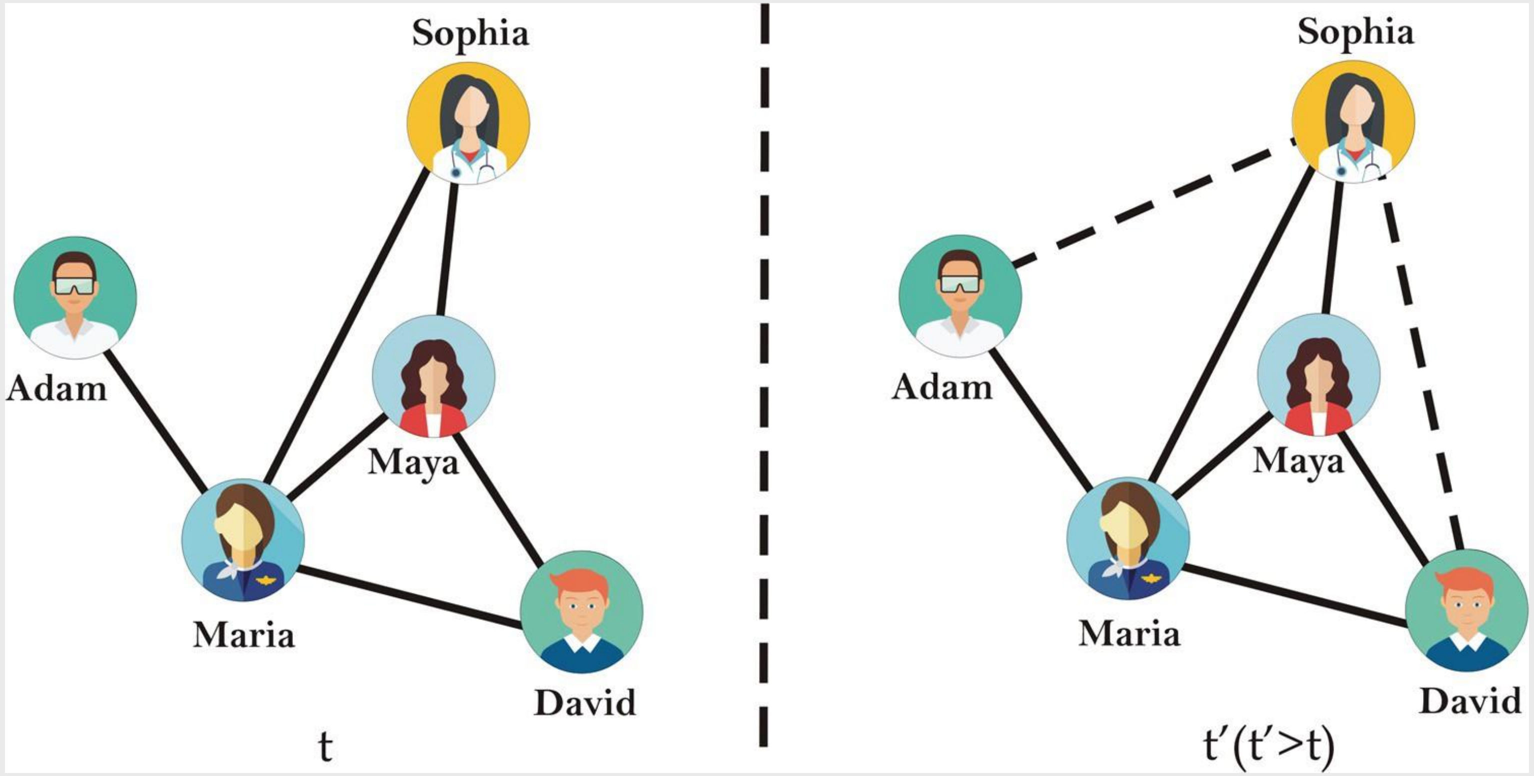
The model uses some assumptions: nodes in the same community have the same type of connectivity pattern

- e.g. there are many links within communities and few links across communities

- **Stochastic Blockmodels** (Harrison White, structural equivalence, core-periphery)
- **Modularity maximization**

More on this on Wednesday

Types of analysis: Link/metadata prediction



Ahmad et al 2020

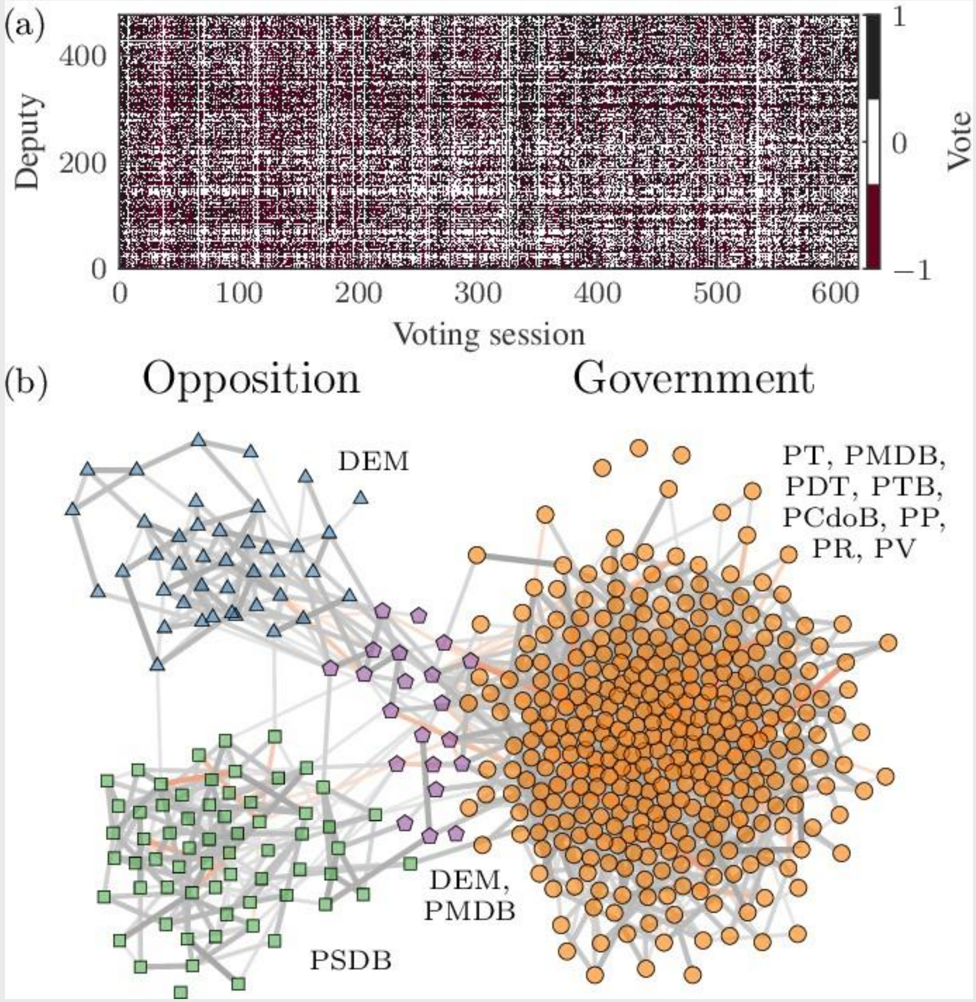
Networks are rarely complete

Approaches such as triangle closure, SBM or node embeddings

More on this on Wednesday

Types of analysis: Network inference

Network from co-occurrences



Network Reconstruction and Community Detection from Dynamics, Peixoto 2019

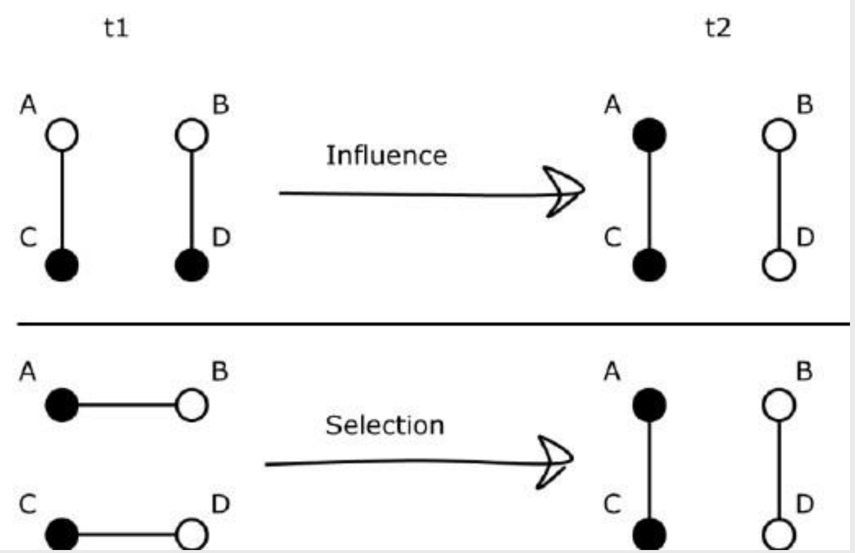
More on this on Thursday

Types of analysis: Dynamics

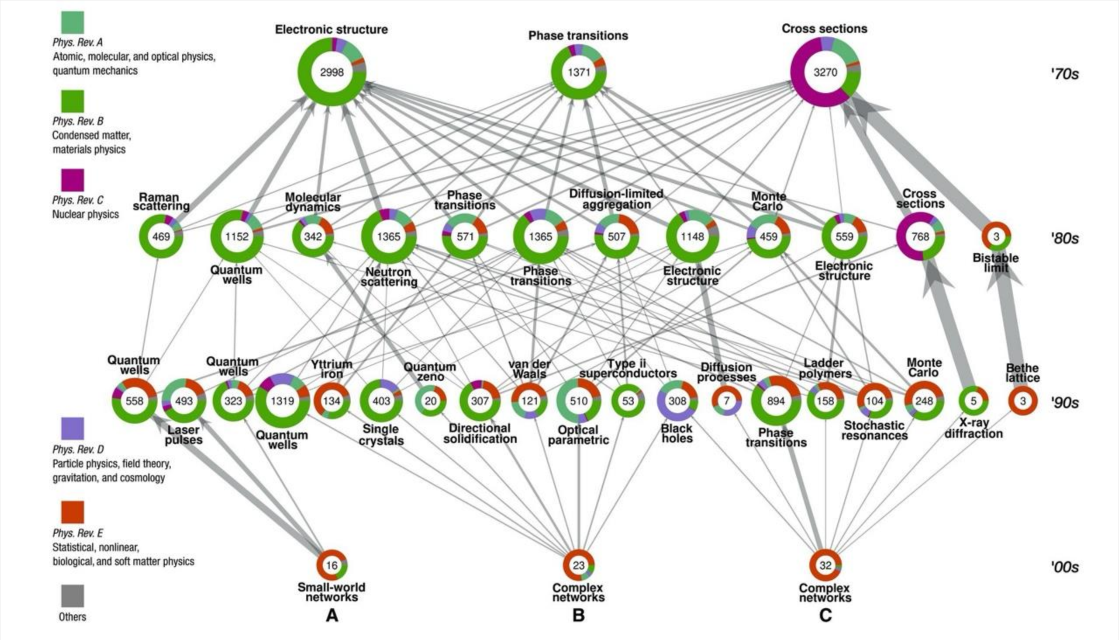
How does behavior/diseases/information spread? Model matters:

Simple contagion vs complex contagion

Longitudinal data: Allow to test selection vs influence, evolution of communities over time, co-evolution of network topology and ideas



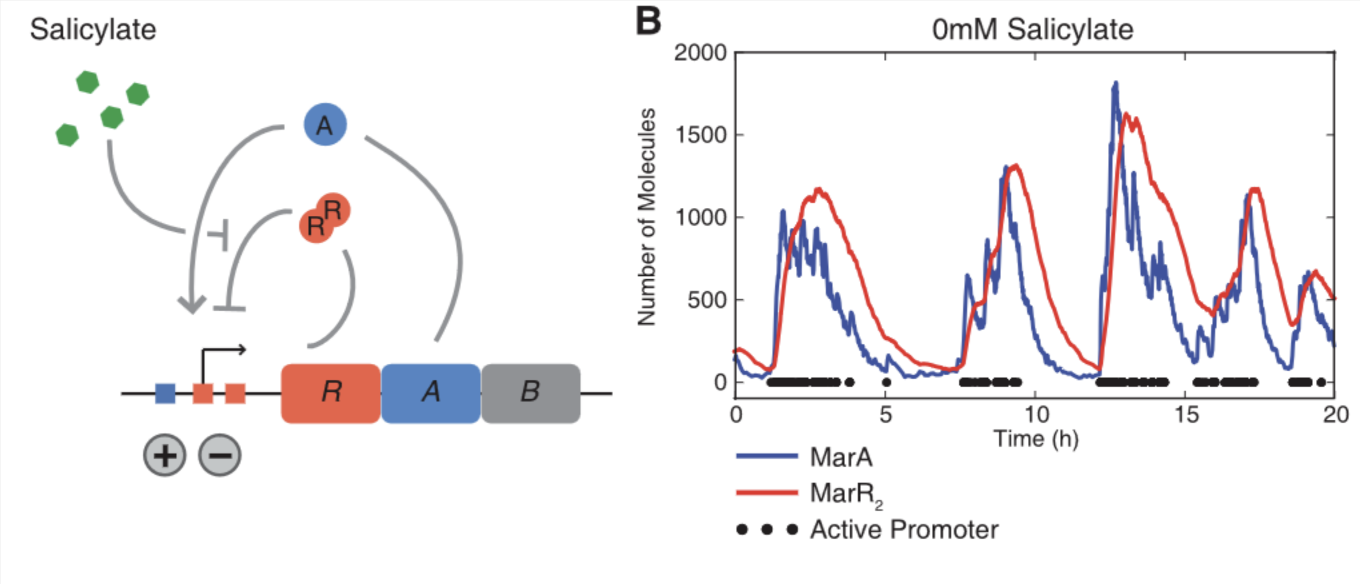
Friemel, 2015



Bovet et al, 2022

More on this on Friday

Types of analysis: Dynamics



$$P_{00}+A \underset{k_{-a}}{\overset{k_a}{\rightleftharpoons}} P_{10}$$

$$P_{01}+A \underset{k_{-a}}{\overset{k_a/\beta}{\rightleftharpoons}} P_{11}$$

$$P_{02}+A \underset{k_{-a}}{\overset{k_a/\beta'}{\rightleftharpoons}} P_{12}$$

$$P_{00}+R_2 \underset{k_{-r}}{\overset{2\cdot k_r}{\rightleftharpoons}} P_{01}$$

$$P_{01}+R_2 \underset{2\cdot k_{-r}}{\overset{k_r}{\rightleftharpoons}} P_{02}$$

$$P_{10}+R_2 \underset{k_{-r}}{\overset{2\cdot k_r/\alpha}{\rightleftharpoons}} P_{11}$$

$$P_{11}+R_2 \underset{2\cdot k_{-r}}{\overset{k_r/\alpha'}{\rightleftharpoons}} P_{12}$$

$$P_{10} \xrightarrow{\alpha_{10}} P_{10}+M+R_{uf}+A_{uf}$$

$$P_{11} \xrightarrow{\alpha_{11}} P_{11}+M+R_{uf}+A_{uf}$$

$$P_{12} \xrightarrow{\alpha_{12}} P_{12}+M+R_{uf}+A_{uf}$$

$$P_{02} \xrightarrow{\alpha_{02}} P_{02}+M+R_{uf}+A_{uf}$$

$$M \xrightarrow{\beta_r} M+R_{uf}$$

$$M \xrightarrow{\beta_a} M+A_{uf}$$

$$A_{uf} \xrightarrow{k_{fa}} A$$

$$R_{uf} \xrightarrow{k_{fr}} R$$

$$2\times R \underset{k_{-dr}}{\overset{k_{dr}}{\rightleftharpoons}} R_2$$

$$R \xrightarrow{\lambda_r} 0$$

$$R_2 \xrightarrow{\lambda_r} 0$$

$$M \xrightarrow{\lambda_M} 0$$

$$A_{uf} \xrightarrow{\lambda_{auf}} 0$$

$$A \xrightarrow{\lambda_a} 0$$

$$R_{uf} \xrightarrow{\lambda_{ruf}} 0$$

$$P_{00} \xrightarrow{\alpha_{00}} P_{00}+M+R_{uf}+A_{uf}$$

$$R_2 \underset{k_{-sal}}{\overset{[Sal]\cdot k_{Sal}}{\rightleftharpoons}} R_2\cdot Sal$$

$$P_{01} \xrightarrow{\alpha_{01}} P_{01}+M+R_{uf}+A_{uf}$$

Garcia-Bernardo and Dunlop (2015,2016, 2017)

Resources

Tools

- Libraries:
- igraph (C, Python & R wrappers)
- Networkx (Pyhon)
- graph-tool (Python (UNIX))
- statnet (R)

- Gephi: open-source network analysis and visualization software package

- Interactive network visualization:
- visNetwork (R) – see e.g. here
- Panel + networkx (Python)
- https://heliosweb.io/docs/example/

Data

Stanford Large Network Dataset Collection:
https://snap.stanford.edu/data/

Network repository:
https://networkrepository.com/networks.php

Netzschleuder:

https://networks.sweked.de

Index of Complex Networks:

https://icon.colorado.edu

Network repository: *networks.sweked.de*

`terrorists_911` — 9-11 terrorist network

Description

Network of individuals and their known social associations, centered around the hijackers that carried out the September 11th, 2001 terrorist attacks. Associations extracted after-the-fact from public data. Metadata labels say which plane a person was on, if any, on 9/11.[1]

1. Description obtained from the ICON project. ↩

Tags

Social Offline Unweighted Metadata

Citation

V. Krebs, "Mapping networks of terrorist cells." Connections 24, 43-52 (2002)., https://doi.org/10.5210/fm.v7i4.941 [@sci-hub]

Upstream URL OK

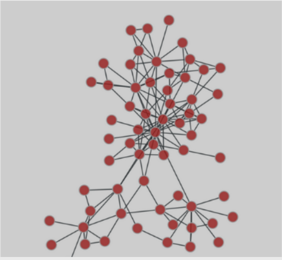
https://aaronclauset.github.io/datacode.htm

Networks

Tip: hover your mouse over a table header to obtain a legend.

Name	Nodes	Edges	$\langle k\rangle$	σ_k	λ_h	τ	r	c	$\oslash$	S	Kind	Mode	NPs
terrorists_911	62	152	4.90	4.00	7.25	19.05	-0.08	0.36	5	1.00	Undirected	Unipartite	id name group

Ridiculograms



Problems with this dataset? Open an issue.
You may also take a look at the source code.

The network in this dataset can be loaded directly from graph-tool with:

```
import graph_tool.all a
g = gt.collection.ns["t
```

`swingers` — Swingers and parties (2013)

Description

A bipartite sexual affiliation network representing "swing unit" couples (one node per couple) and the parties they attended.[1]

1. Description obtained from the ICON project. ↩

Tags

Social Offline Unweighted

Citation

A.-M. Niekampab et al., "A sexual affiliation network of swingers, heterosexuals practicing risk behaviours that potentiate the spread of sexually transmitted infections: A two-mode approach." Social Networks 35(2), 223-236 (2013), https://doi.org/10.1016/j.socnet.2013.02.006 [@sci-hub]

Upstream URL 404

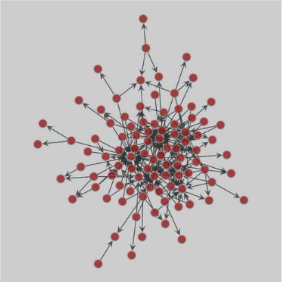
https://sites.google.com/site/ucinetsoftware/datasets/covert-networks/swingers

Networks

Tip: hover your mouse over a table header to obtain a legend.

Name	Nodes	Edges	$\langle k\rangle$	σ_k	λ_h	τ	r	c	$\oslash$	S	Kind	Mode	NPs	EPs
swingers	96	232	2.42	5.19	7.46	5.19	-0.34	0.00	7	1.00	Directed	Bipartite	name	2

Ridiculograms



Problems with this dataset? Open an issue.
You may also take a look at the source code.

The network in this dataset can be loaded directly from graph-tool with:

```
import graph_tool.all a
g = gt.collection.ns["s
```

Practical 1, exercise 4

Recap

There is important information encoded in relationships/interactions

Modeling systems using networks allow us to study that information

We can represent networks using adjacencies matrixes or adjacencies lists

We can **describe networks**: number of edges and nodes, components, density, assortativity, clustering, diameter, degree distributions.

We can **test hypothesis** using network models.

There are many analysis:

- We can find the most important nodes using **centrality measures** -- Today
- We can test hypothesis using **statistical models** -- Tuesday
- We can find clusters of nodes using **community detection algorithms** – Wednesday
- We can **predict links and metadata** -- Wednesday
- We can reconstruct the network using **network inference** -- Thursday
- We can analyze **dynamics** (e.g. contagion) -- Friday