NME 2024

Network Modeling for Epidemics

1

Start with some motivation

Why do we need this network stuff?

What are the other frameworks missing?

- DCMs and ICMs have difficulty representing complex partnership networks
  - Heterogeneity in mixing patterns across an arbitrary number of attributes
  - Cross-partnership influences on formation and dissolution processes

- And they lack a ***general*** framework for estimating partnership network patterns from data
  - As a result, they are often limited to one or two features
  - Adding more ranges from very difficult to impossible

- So assumptions often replace data and empirical assessment

We make lots of assumptions in epidemic models

- The goal is not to eliminate assumptions
  - It's not to make them blindly
  - And not to let the limitations of a framework dictate them
- We should try to understand the impact of assumptions
  - On the transmission system
  - And the epidemic outcomes of interest
- And choose a modeling framework that captures the key impacts

Not this: Framework → Assumptions

But this: Assumptions → Framework

So: Do detailed partnership patterns matter?

- Yes, in some contexts
  - When contact is needed for transmission
    - E.g., respiratory infections vs. water-borne infections
  - "Effective contacts" are relatively rare, so the network is "sparse"
    - E.g., sexually transmitted infections vs. respiratory infections
  - And duration of at least some partnerships is relatively long

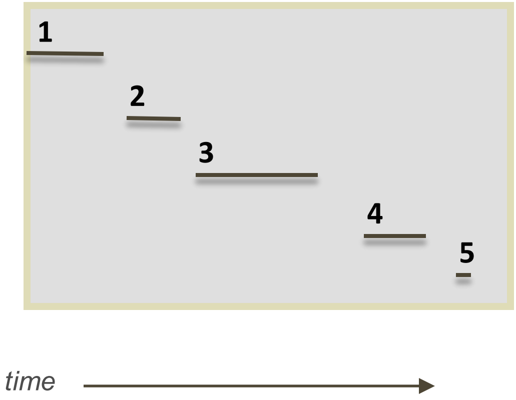
- Then detailed partnership patterns determine the connectivity of the transmission network

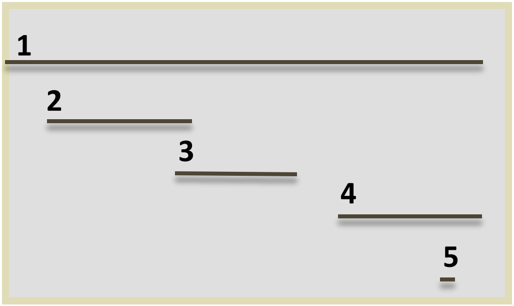
- In two ways
  - Cross-sectionally (at a moment in time)
  - Longitudinally

Example: Partnership timing and sequence

A sequence of partnerships can be either :

Serially monogamous *or* Concurrent

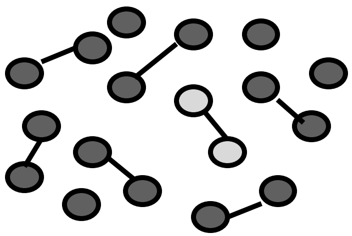




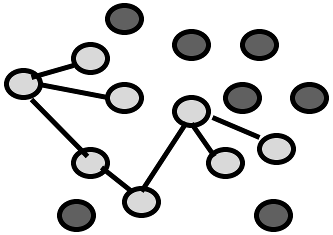
Same number of partnerships,
but the sequence of start and end dates is different

Concurrency and cross-sectional connectivity

A unique *cross-sectional* network signature:



monogamy



concurrency

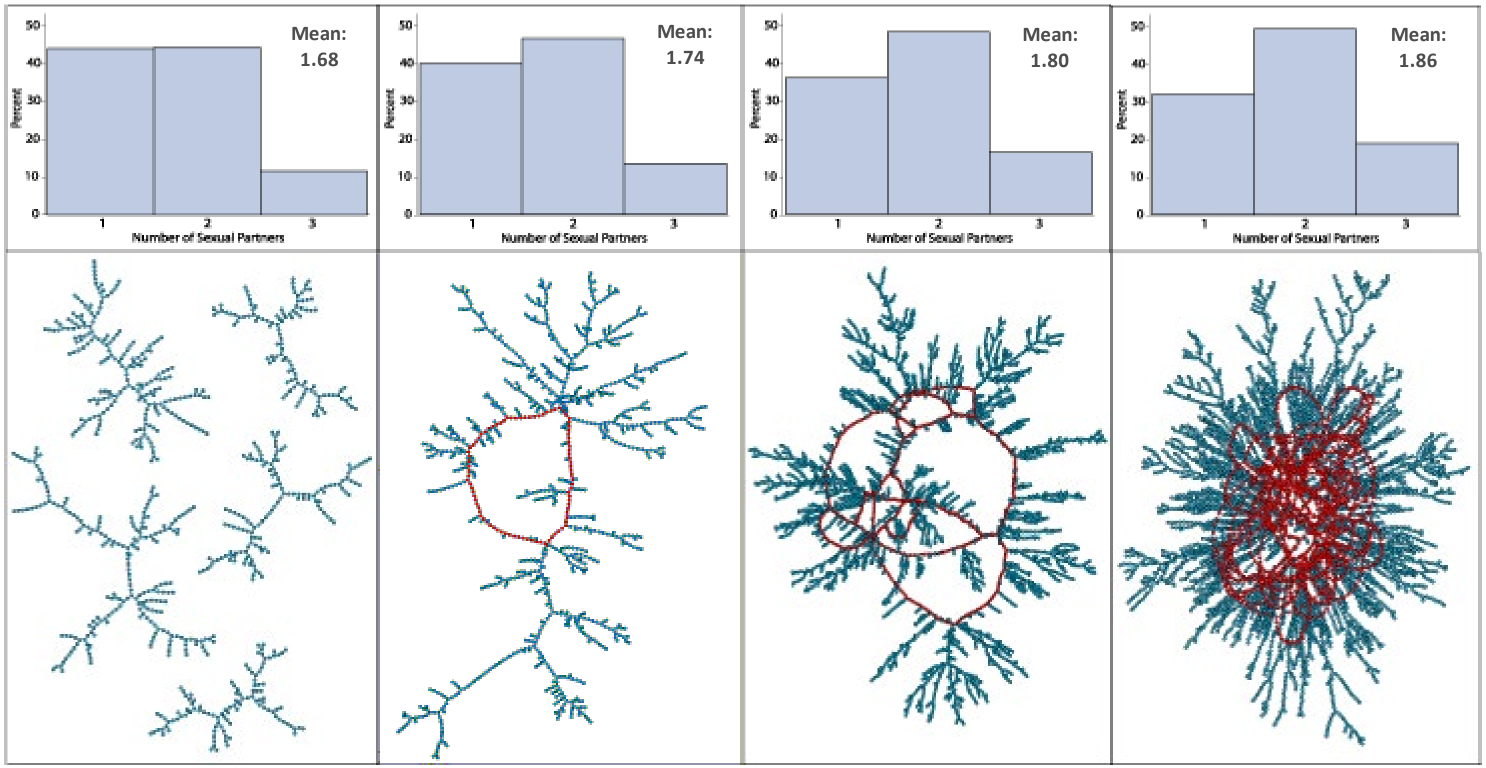
Note: even with p(transmission=1), and duration=forever, there would be no spread in this network unless partnerships are dynamic

Concurrency and cross-sectional connectivity

Number of ongoing partners on any particular day

Largest components

Bicomponents in red

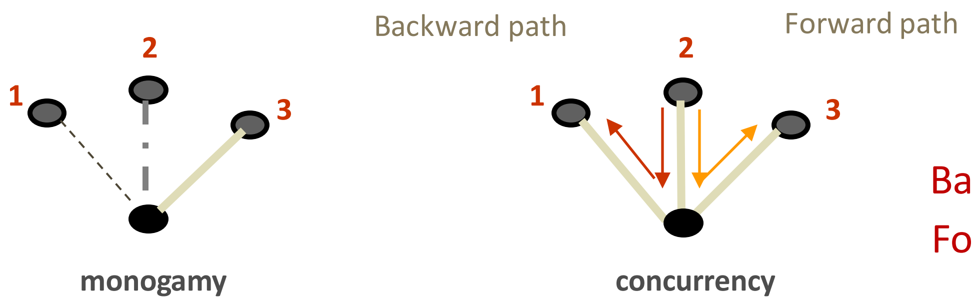


In largest component:	2%	10%	41%	64%
In largest bicomponent:	0	1%	5%	15%

Ref: Morris, Goodreau and Moody 2007

Concurrency and longitudinal connectivity

Concurrency removes the protection of sequence over time



Backward path: New chain of infection

Forward path: Shorter generation interval

This changes two features of epidemic dynamics:

The reachable path of infection

Epidemic velocity

The reachable path

- The *reachable path* in a network is determined
  - Not by the partnerships at any single point in time
  - Nor by the cumulative total over time

- But by the

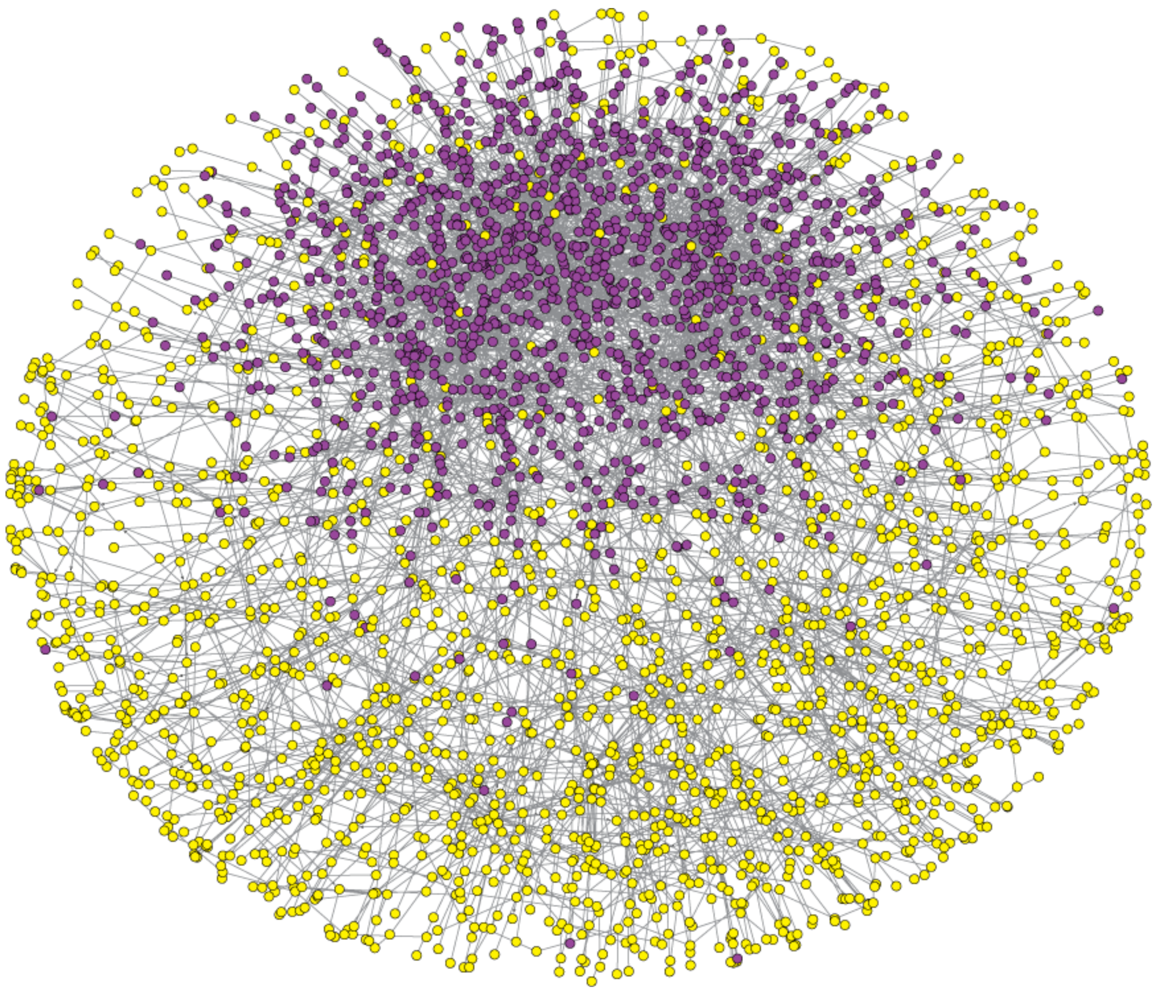
  cumulative time ordered sequence of partnerships

Cumulative connectivity: a network collapsed over time

10,000 nodes

2 demographic groups

10 years
of partnership
activity



99.7%

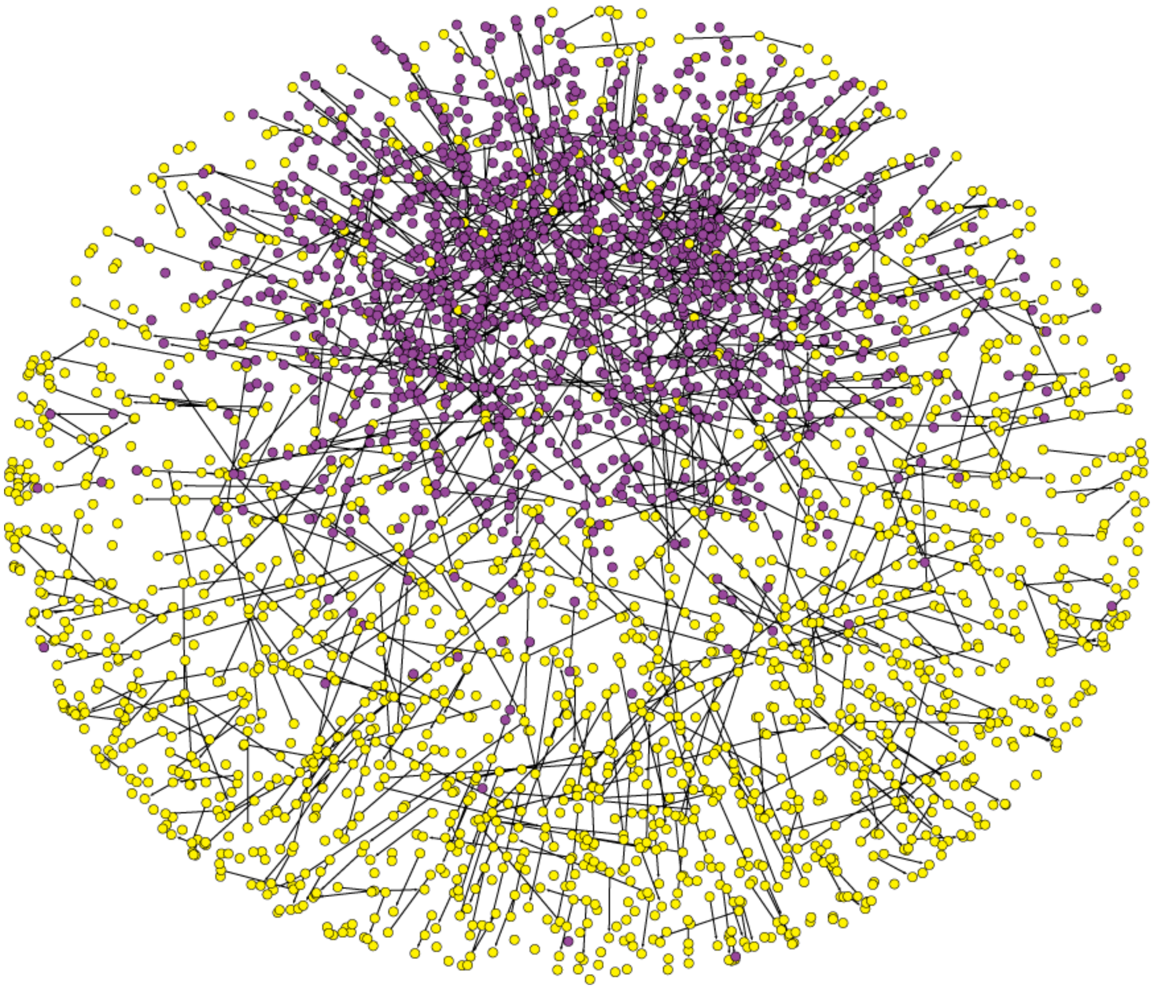
Connected

(nodes that can
reach each other
directly, or
indirectly)

Daily connectivity: A momentary snapshot

Almost all components are size 2 or smaller

The largest components have 5-6 nodes



0.06%

connected

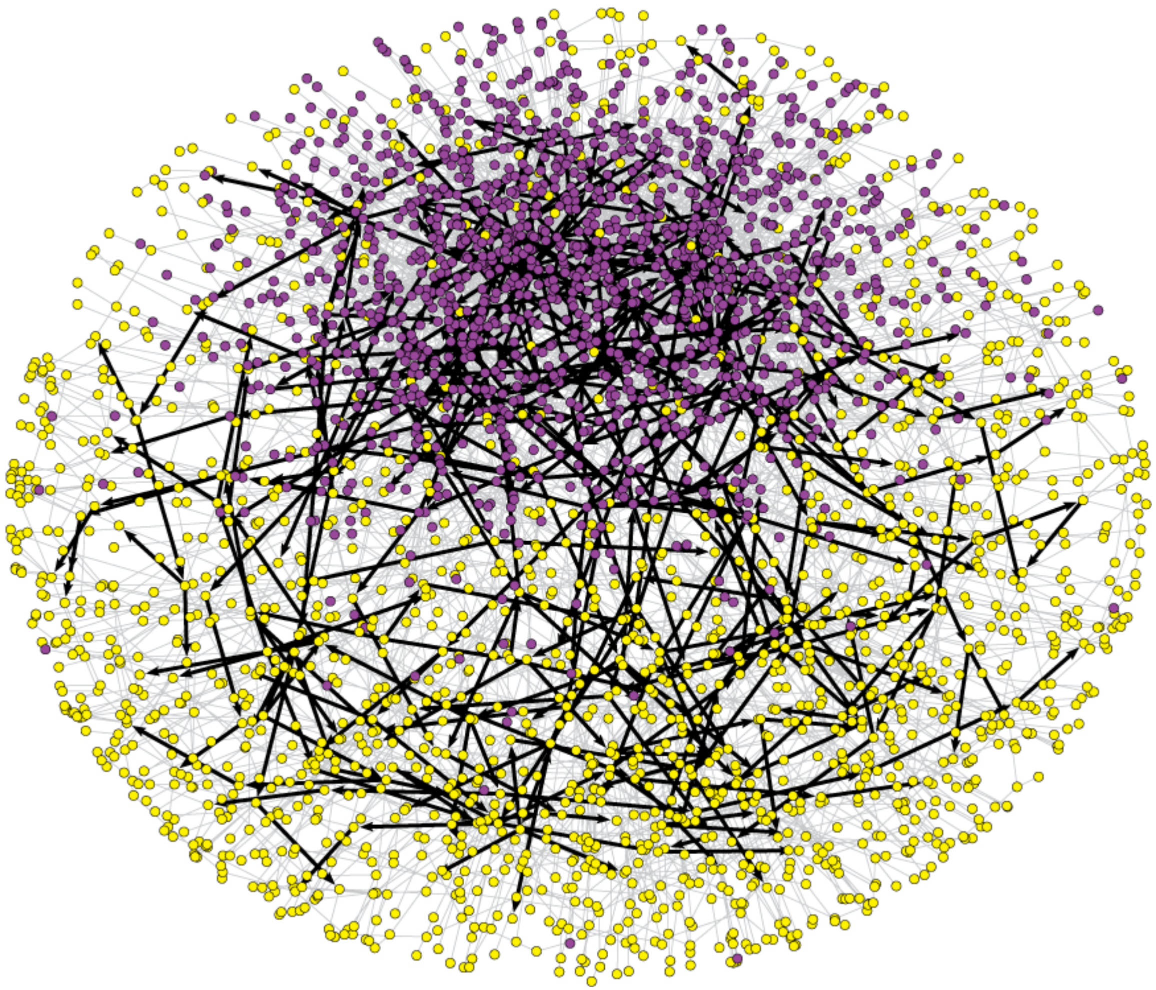
The reachable path: different than both of these

From 10 initial seeds

Trace out:

The cumulative time-ordered path of partnerships

Over 10 years



5.0%

Reachable

Track the growth of the forward reachable path (FRP)

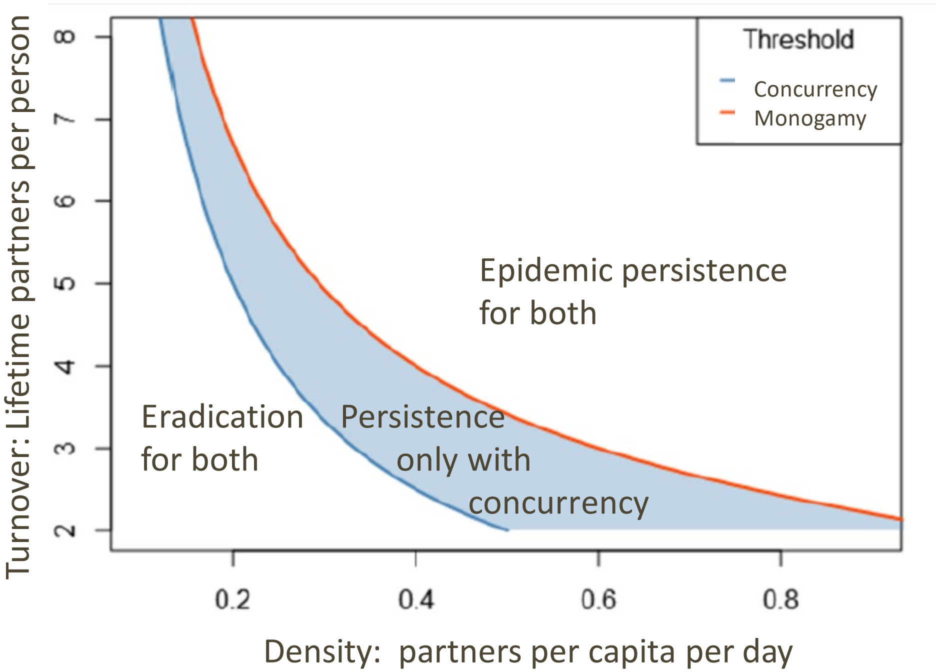
- Original network: 10,000 nodes
- Focus here on the ~600 that end up in the FRP
- 10 initial seeds (slightly larger squares)
- Tie color indicates:
  - BLUE: join FRP when dyad is mutually monogamous
  - RED: join FRP when at least one node has concurrent partner
  - BLACK: active tie, but dyad not in FRP yet

Only **5%** of ties are ever concurrent

But these account for **> 50%** of the FRP

Movie online at:
https://www.youtube.com/watch?v=r3LYA5kirjA&t=214s

Concurrency and R_0



- Epidemic persistence is a function of both static and temporal connectivity
  - **Density**: partners per capita on any day
  - **Turnover**: partners per capita over the lifetime
- Concurrency = more connectivity
  - Lowers the threshold for persistence
  - For the same values of density and turnover
- So there is a region where epidemics persist only if there is concurrency

Armbruster et al. (2017)

So yes, persistent partnerships matter

- For the connectivity of a transmission network
  - Cross-sectionally (with clear thresholds)
  - Longitudinally – via the reachable path and velocity

- And this is why we need network models
  - To specify the patterns & timing of partnerships in a network
  - And to be able to estimate these, in a principled way, from data

Terminology: what is a network model?

- The word "model" is used in many ways in this setting
  - The overall epidemic model
  - Submodels for components of the overall epidemic model
- Epidemic models can contain many submodels
  - To control different processes – demographics, disease progression, pathogen life cycles, etc.
- So for clarity, in this part of the course:
  - "Network epidemic model" – an overall epidemic model that represents the transmission network explicitly
  - "Network model" – the submodel for *just* the network

What is a good network model?

A dedicated model for just the network process:

- With a general framework
  - Not restricted to a single type, like scale-free, small world, or age mixing
  - Able to represent an arbitrarily wide range of properties

- And a principled statistical foundation
  - A stochastic model for the probability of the network
  - With methods for estimating from data
  - And methods for model assessment

A general statistical model

- Model: LHS = RHS
  - Something like: outcome = f(covariates)
    - For example: income = f(age, sex, race, occupation, …)

- A general statistical model allows you to
    - Specify a model for an outcome from a (general) class
    - Estimate the impact of one or more covariates from sampled data
    - Conduct statistical inference (assess models and quantify uncertainty)
    - Predict and simulate outcomes from a fitted model
    - For example: OLS $\hat{y} = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 sex + \beta_4 race$

A general statistical network model

- LHS = RHS, so here: Network = f(covariates)
  - What covariates might you find in a network model?
    - What influences the systematic structure of a network?
    - What about the dynamics? (formation & dissolution rates)

- A general statistical network model allows you to
  - Specify network models from a broad, general class
  - Jointly estimate the impact of covariates on the structure & dynamics of a network, from (possibly sampled) data
  - Conduct principled model assessment and selection
  - And predict / simulate networks from the model

NME 2024

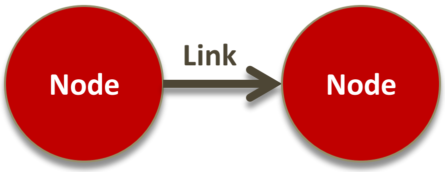
Network Modeling for Epidemics

Network concepts and terminology

Some basics for setting up the rest of the week

Terminology

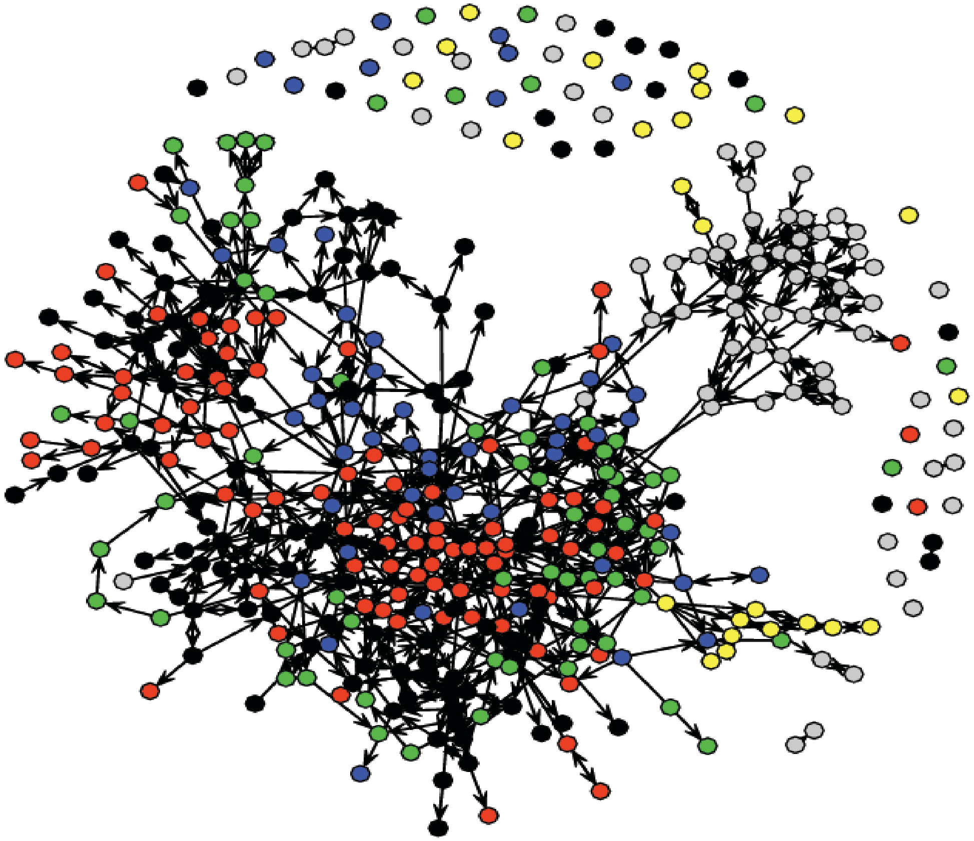
- **Node**: the entity of interest
  - for us, nodes represent people; also called actors or vertices
- **Link**: the relationship of interest
  - also called a tie, an edge, or a line



- **Network**: a collection of nodes and links
  - also called a graph

Nodes, links and networks

Beyond the pretty pictures, there are many different attributes of nodes, links and networks that have implications for the structures we can observe, and what we want to model



Types of nodes

- Individual units
  - Humans
  - Animals
  - Airports
  - Computers
  - Genes

- Collectivities
  - Countries, cities
  - Families
  - Species
  - Organs, Sensory systems

In social networks, a focal node is called “ego”, and the nodes linked to this focal node are “alters”

Nodes have lots of properties we call “attributes” (age, size, etc.)

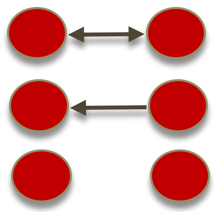
Types of links

- Social
  - Affective (like/dislike, trust/do not trust)
  - Kinship / social role (mother of, brother of, boss of)
  - Exchange (advice seeking, sexual intercourse, trade)
  - Cognitive (knows/does not know)
  - Affiliation (belongs to, is a member of)
- Physical
  - Road
  - Flight path
  - Wire / Wireless

Links can also have attributes: length, type, etc.

Link properties

- Directed (e.g., likes)
  - Mutual
  - Asymmetric
  - Null



Nodes are now classified as senders and/or receivers

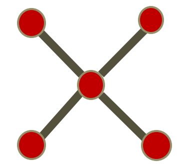
  - A directed graph is also called a di-graph
  - A directed edge is also called an arc
- Undirected (e.g., has sex with)

- Binary (0,1 on or off only)
- Signed and/or Valued (… -2, -1, 0, 1, 2 …; or continuous values)

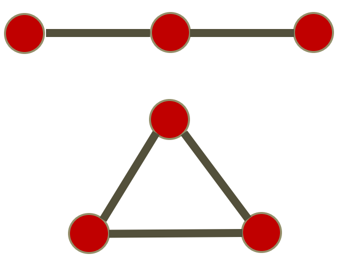
Configurations / Subgraphs

Dyads

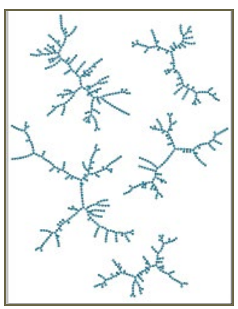
Stars



Triads



Components



Any collection of nodes and links can be defined as a configuration

Levels of measurement

As we look at ways of describing network data, keep in mind the different levels of measurement

- Node level: *attributes of individual nodes*
  - Examples: age, sex, infection state, degree
- Dyad level: *attributes of pairs or links*
  - Examples: type of relationship, duration
- Component level: *subgraph attributes and distributions*
  - Examples: size, density, degree and geodesic distributions …
- Network level: *overall structural attributes and distributions*
  - Examples: density, degree, geodesic distribution …

E.g. Cycles

- Paths that lead back to the origin node
  - Cycle length k = number of lines in the cycle, "k-cycles"
  - Triangles are 3-cycles

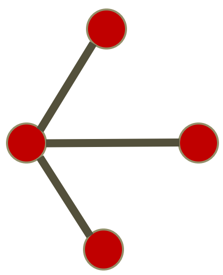
- Node level measure: Number of cycles a node is a member of

- Edge level measure: Number of cycles an edge is a member of

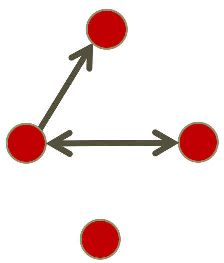
- Network level measure: The "cycle census", the frequency of these configurations in the network

Common network level measures

- Density: Fraction of all dyads that have an edge
- Isolate count: Number of nodes without any edges



Nodes: 4
Isolates: 0
Dyads: (4*3)/2 = 6 (undirected)
Density: 3/6 = .5



Nodes: 4
Isolates: 1
Dyads: 4*3 = 12 (directed)
Density: 3/12 = .25

Types of networks

- Simplest: 1-mode, undirected, binary ties, single relation

- 2-mode (aka *Bipartite*)
  - Two different types of nodes
  - Ties only allowed between groups

  Examples: Online network groups and persons (Affiliation network); *heterosexual sex net*

- Multiplex
  - More than one type of link possible

  Example: Workplace and home

- Multilevel
  - Hierarchical/nested or Overlapping

  Example: Grades within schools; persons working across multiple projects

Representing network data

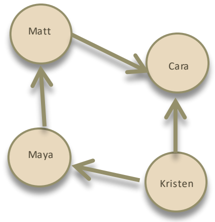
- Sociomatrix
  - (aka adjacency matrix)
  - simple but inefficient for large sparse nets (order n^2)

	Matt	Cara	Kristen	Maya
Matt	0	1	0	0
Cara	1	0	0	1
Kristen	0	1	0	1
Maya	1	0	0	0

- Edgelist

Matt	Cara
Cara	Matt
Cara	Maya
Kristen	Cara
Kristen	Maya
Maya	Matt

- Graph



What’s next?

- *Take a break now*

- *When we come back: A short lab*
  - To get some hands on practice with descriptive network analysis
  - Using `statnetWeb`