Network Models

Questions From the Last Lecture??

Compartmental Model Recap

- Population is divided up into compartments based on disease state/demographics
- Population mixes randomly
- Beta is a composite of contact and probability of infection given contact, and conventionally non-identifiable

Networks as a Concept

- What is a “network”
  - A conceptual way of representing relationships between things
  - Nodes: *Things*. People, places, etc.
  - Edges: Links between nodes
  - Occasionally these are called graphs, vertexes and arcs
    - Network science co-evolved in several different fields at about the same time
- Networks can be represented in a number of different ways
- A network’s structure is sometimes called its “topology”

Representing Networks

Diagram

1	2
2	3
2	4
3	4

Edge List

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Adjacency Matrix

Representing Networks

- Diagrams
  - Pros: Easily visualize the network structure, often look really cool
  - Cons: Can get difficult to interact with rapidly, “hairball” networks, not easily machine readable
- Edge Lists
  - Pros: Compact, expressive, easily machine readable
  - Cons: Less “human readable”
- Adjacency Matrix:
  - Pros: Matrix operations unlock all kinds of cool analysis techniques
  - Cons: Also less human reasable, less machine readable than edge lists
- Easy to go back and forth

Complications

- Directed Networks

- Weighted Networks

- Time-varying Networks

- Different Types of Nodes/Edges

Very Complex Networks

- Multi-scale Networks
  - Networks of Networks
- A network can simultaneously be directed, weighted, time-varying, etc.

Many Networks, Many Questions

- Network Analysis is interested in the study of the shape and structure of networks
  - How are networks connected, do they form patterns, etc.
  - Simulating a process (like an epidemic) on a network is a small part of that
- Very active research field – sociology, physics, computer science and math all have their own takes

Basic Network Terminology

- Size: The number of nodes
- Density: The ratio of the number of edges in a network compared to the number of *possible* edges
- Path: A set of edges connecting node *I* and node *J*
- Shortest Path: A path with the least number of edges

More Terminology!

- Complete Graph: A network where all nodes are connected
- Connected Component: A portion of a network where there is a path between every pair of nodes
- Clique: Subsets of a network where every node in the clique forms a complete sub-graph
- Clustering Coefficient: The ratio number of neighbors of a node who are connected to each other to to the maximum possible number of such connections

Centrality

- Ways of describing a node *i*'s position in the network
  - Degree: How many other individuals connect to *i*?
    - In-degree: In a directed network, how many incoming connections
    - Out-degree: In a directed network, how many outgoing connections
  - Betweenness: How many shortest paths between individuals go through *i*?
  - Closeness: How many contacts away is *i* from all other individuals?
  - PageRank: How likely are you to end up at a particular node given you're wandering along paths with the occasional random jump. This is how Google got started.

Types of Networks

- Generated/Theoretical
- Empirical

Erdős–Rényi random graph:

- Edges form between nodes with probability *p*

Watts-Strogatz Small World Network

- Each node is connected to its *k* closest neighbors. There is another probability, *p,* that each edge will be “rewired”, and instead become a random connection
- This results in a highly clustered network with occasional links between communities
- This is the “Six Degrees of Separation” concept

Barabási–Albert Preferential Attachment

- "Rich get richer"
- Slowly grow the network one node at a time
- Each node *preferentially attaches* to an existing node
- This results in a degree distribution that follows a power law
- Resembles some real systems, how *closely* it represents some real systems remains unclear
- Can result in interesting results – on average, if you pick someone at random, their friends have more friends than they do

$$p_i = \frac{k_i}{\sum_j k_j}$$

Empirical Networks

- These are networks that are collected from the real world
- These are often very useful, but are also often *very* expensive in terms of the time, effort and cost to collect them
- Extensive concerns about privacy
- *Network Epidemiology* by Martina Morris discusses the process behind collecting many networks

Surveys

- Ask people who they know
- Often simple and straightforward, but may vary on memory, how people define “friends”, etc.
- FluScale study in China is an example of this
- How do you reach vulnerable populations?
- Can build a network off each persons answers, or try to recruit those answers (”snowball” or respondant driven sampling)

Colocation

- Assume people in the same place at the same time know each other
- Actors in the same movie, socialites at the same parties, people in the same work environment

Scrape from Data

- Hospital transfer information
- Airline networks
- Livestock transactions
- Can often be expensive/proprietary

Sensors

- Can automatically collect who is near each other using sensors, mobile phones, etc.
- Assumes nearness = contact
- May be true for some diseases (influenza) and not for others (STIs)

How Are Network Models Implemented?

- Proving things about networks is hard, and beyond the scope of this class
- Often done by simulation

A Simple Network Model

- Start with one infected individual

A Simple Network Model

- Spreads along edges to other nodes at probability p, where $p = 0.65$

A Simple Network Model

- Spreads along edges to other nodes at probability p, where $p = 0.65$

A Simple Network Model

- Spreads along edges to other nodes at probability p, where $p = 0.65$

A Simple Network Model

- Spreads along edges to other nodes at probability p, where $p = 0.65$

A Simple Network Model

- Spreads along edges to other nodes at probability p, where $p = 0.65$

A Simple Network Model

- Spreads along edges to other nodes at probability p, where $p = 0.65$

Differences Between Compartmental and Network Models

- Results make not agree
- This can be the result of:
  - Random chance (network models are by-and-large stochastic)
  - Network structure
- Contact pattern – note that each node has fewer interactions than they would in a compartmental model

Network Strengths

- More direct modeling of contact
- Can decouple contact and probability of infection transmission given contact
  - These two are nonidentifiable in a compartmental model
- They look really cool
- Built-in ability to have heterogeneity, stochasticity, etc.

Network Drawbacks

- Where does your network come from?
  - Some networks can be assumed
  - Likely better than mass action, but still possibly not right
- Collecting empirical networks is hard
- Much higher computational requirements
- Much harder to prove things
- Largely reproducible given the network, but many people don’t make their networks public