Processes Taking Place on Networks

- Traffic flow on roads
- Airline flights between airports
- Data traffic on the Internet
- Reactions in metabolic networks
- Energy flow in food webs
- Spread of computer viruses between computers
- Spread of news or rumors over social networks
- Spread of diseases between people
“Expander” property

• A surprising result is that most networks possess the “expander” property:
  − The average surface area of a subset of the graph is proportional to the subset's volume

• Our common experience with things that are not networks is that surface area is usually “smaller” (in some sense) than volume
And in general dimension $d$

$$A \propto V^{(d-1)/d}$$

$$A = 4\pi r^2$$

$$V = \frac{4}{3} \pi r^3$$

$$A = \sqrt[3]{36\pi} V^{2/3}$$
• But in a network the surface can be proportional to volume
• Equivalent to choosing $d$ infinite
• We gain a constant *fraction* in volume every time we expand our region by one step

• Call this fraction $z$. After $s$ steps we have reached at least $z^s$ nodes

• It will take around $s$ steps to cover all nodes, where $n = z^s$, or

\[ s = \frac{\log n}{\log z} \]
Example: A computer virus
FIGURE 2. Probable cases of severe acute respiratory syndrome, by reported source of infection* — Singapore, February 25–April 30, 2003

* Patient 1 represents Case 1; Patient 6, Case 2; Patient 35, Case 3; Patient 130, Case 4; and Patient 127, Case 5. Excludes 22 cases with either no or poorly defined direct contacts or who were cases translocated to Singapore and the seven contacts of one of these cases.

Epidemiology

• As a disease spreads through a population, individuals pass through different states

• In the simplest case there are two states:
  - Susceptible (S) means you don't have the disease
  - Infective (I) means you do and you can pass it on

\[
\begin{array}{c}
S \\
\rightarrow
\end{array}
\begin{array}{c}
I
\end{array}
\]
• We can add more states:
  − Recovered (R) means you have had the disease, gotten over it, and *are now immune*
  − Exposed (E) means you have caught the disease, but are *not yet infective* (though you will be)
• We can also change the flow-chart
  – Individuals who have recovered can lose their immunity and become susceptible again:
Traditional approaches

- Traditionally the way to model diseases was to make the full mixing assumption
  - A susceptible individual catches the disease from an infective one
  - The probability of transmission between any two individuals in given compartments is the same for all individuals in those compartments
- This is obviously wrong, but it's convenient because it makes the math easier
Traditional approaches

• For example: the simple SIR model is then governed by the equations:

\[
\frac{dS}{dt} = -\beta \frac{SI}{n},
\]

\[
\frac{dI}{dt} = \beta \frac{SI}{n} - \gamma I,
\]

\[
\frac{dR}{dt} = \gamma I.
\]
Network approaches

- Once we bring networks into the picture, the differential equation approach breaks down. We need other techniques.

- One crucial observation is that diseases are more likely to find people with many connections:
Disease spreads if \( R_0 > 1 \) or if

\[
T > \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}
\]
• For networks with highly skewed degree distributions, \( \langle k^2 \rangle \) can be very large, making \( T \) very small

• For power-law degree distributions \( \langle k^2 \rangle \) can diverge in the limit of large network size, meaning \( T = 0 \) and the disease *always* spreads

• We need to be careful however. This result can be modified by other things such as
  - Correlations in the network
  - Local network structure

• And there are many other things we can calculate. . .
We studied an outbreak of Mycoplasma pneumonia in a hospital in Evansville, IN

- 15 wards, 250 patients, 440 employees
- 60 patients and 82 employees diagnosed with *M. pneumoniae*
- Patients are confined to wards but caregivers move between them
- Standard control strategies are patient cohorting and chemoprophylaxis for patients (Azithromycin)

Control strategies

- Cohorting ineffective: long incubation period
- Caregivers primary vectors, even though they don't get sick
- Essentially 100% transmission from infected caregivers to patients

Recommendations:
- Limit caregivers to a single ward
- Give antibiotics to caregivers
- Other strategies to prevent caregivers being infected
Simulated contact network

Incorporates real data from the city of Vancouver, BC on:

- Sizes of households
- Age distributions
- Working population
- School-age population
- Sizes of school classes
- Hospitals, wards, healthcare workers, beds
- etc.
Sexual contact networks

Foxman, MEJN, Percha, Holmes, and Aral,  

- Relatively precisely defined example of contact network
  - People can usually remember who they slept with
  - Reasonably clear definition of what constitutes a contact

- But:
  - Network not static, changes on short timescales
Survey data

- Random-digit dialing survey:
  - 37,000 phone numbers called in Seattle, WA area
  - 8683 answered and were non-commercial
  - 2582 met eligibility criteria, of which
  - 1051 agreed to participate and had ever had sex

- Computer-guided survey asked about 5 most recent partnerships:
  - start and end dates, frequency and type of contact
  - demographics and some personal data on interviewee
Some summary statistics

- Median length of partnership: **92.2 days**
- **38%** had 5 partnerships or fewer in their lifetime
- **15%** had 20 or more partnerships
- Median gap between partnerships: **121.6 days**
- Fraction of partnerships that *overlap* with the preceding partnership (concurrency): **25.6%**
- Median length of overlap: **427.3 days**
- New partnerships: **8.1 per year** per single person
What does it mean?

- 59% of gaps are less than six months, i.e., within the likely infectious period of many STIs

- 33% of gaps are less than eight weeks, the typical infectious period of HIV

  - Thus even non-concurrent relationships are effectively concurrent

  - The structure of the network changes substantially on timescales relevant for the spread of HIV and other diseases, implying that simple static network models are inadequate to predict epidemiological outcomes
• Time ordering of relationships is important for epidemiology: a person cannot be infected by a partner who will acquire HIV in the future.

• Surprisingly few correlations were found with demographic parameters such as age and income.

• Lifetime numbers of partners has a fat-tailed distribution, but a simple network model with a fat-tailed degree distribution is nonetheless inappropriate.
Acknowledgments

• Collaborators:
  - Betsy Foxman, Beth Percha (Michigan)
  - Lauren Meyers (UT Austin)
  - Babak Pourbohloul, Bob Brunham (UBC)
  - Mike Martin, Stephanie Schrag, Sevgi Aral (CDC)

• Funding:
  - National Science Foundation
  - McDonnell Foundation