Co-evolving Complex Networks: Epidemics in Social and Wireless Networks

Madhav V. Marathe

Dept. of Computer Science & Network Dynamics and Simulation Science Laboratory
Virginia Bioinformatics Institute
Virginia Tech
NDSSL TR-09-014
These slides are a slightly extended version of the invited presentation given on March 02 2009 at Miami USA. A few additional slides are included to make the talk more self contained. Additionally, the last few slides contain a list of references that readers might find useful.


SIAM Annual Conf on Computational Science and Engineering (CSE’09)

Venue & Date: Miami March 02 2009.
**Acknowledgements**

Virginia Tech: Network Dynamics & Simulation Science Laboratory, VBI, our colleagues at Los Alamos, Aravind Srinivasan, Tom Dubois (UMD)

Work funded in part by NIGMS, NIH MIDAS program, CDC, Center of Excellence in Medical Informatics, DTRA CNIMS, NSF, NeTs, NECO and OCI program, VT Foundation.
Preparing for pandemics

- **1918 Pandemic**
  - 50 million deaths in 2 years (3-6% world pop)
  - Every country and community was affected

- **Good news**
  - Pandemic of 1918 lethality is currently unlikely
  - Governments better prepared and coordinated: e.g. SARS epidemic

- **But ..**
  - Planning and responding to even a moderate outbreak is challenging:
    - inadequate vaccines/anti-virals, unknown efficacy, hard logistics issues
  - Modern trends further complicate planning:
    - increased travel, immuno-compromised populations, increased urbanization
Models in Mathematical/Computational Epidemiology

- **Mathematical Models for Epidemiology**
  - **Differential Equation Based**
    - [Hethcote: SIAM Review]
    - ODE's
      - [Ross, McDonald, Hamer, Kermack, McKendrick]
    - Stochastic ODE's
      - [Bartlett, Bailey, Brauer, Castillo-Chavez]
  - **Network-Based Modeling**
    - [Newman SIAM Review]
    - Simple Random networks
      - [Barabasi, Moore, Newman, Meyer, Vespignani]
    - Realistic Social Networks
      - [Eubank et al., Marathe, Longini et al., Ferguson et al.]

Network Dynamics and Simulation Science Laboratory
1. Create a synthetic population
   • Sampling Contingency Tables, Assignment Problems

2. Derive a social contact network $G$
   • Assign activities (CART Trees), locations (Gravity models), Construction and analysis of large networks

3. Create a model of disease transmission
   • Design probabilistic timed finite state automata based on data

4. Simulate disease spreads over $G$
   • Simulation of a diffusion process

5. Study effect of interventions: co-evolution of $G$, behavior, policy and disease progression
   • Markov decision processes (MDP) and $n$-way games

Practical Usefulness

- White House Homeland Security Council for smallpox mass vaccination
  - Do we need mass vaccination? How do we protect critical workers?

- Top-Off 2 outcome analysis
  - Socio-economic analysis of interventions

- Multi-sector disruption -- NISAC DHS Study
  - Situational awareness and coordination with multiple infrastructures

- Federal Influenza Plan: OHS & DHHS -- NIH MIDAS project
  - TLC: Targeted Layered Containment, Importance of Social Distancing

- Pandemic Planning for National Guard Preparedness: DoD
  - Impact of layered interventions for force projection: Public versus military health epidemiology

- Pandemic Preparedness for Medcom: DoD
  - Can we develop general guidelines for military populations?

- DHHS Medkit study
  - Use of markets in conjunction with public stockpiles to distribute A/V
Difficulty: Very Large Co-evolving Networks

- Network has to be synthesized: Data is sparse, incommensurate
  - Need new methods for information fusion: Currently using 34 databases

- Large Complex Large Networks
  - >100GB input data: 300M people, 22Billion edges, 100M locations, 1.5B daily activities
  - Irregular Network: Dimension reduction techniques do not apply
  - Co-evolving behaviors, networks and disease spread

- Large experimental design ⇒ multiple configurations
  - 5000 run study not unusual

Unique challenge for researchers in HPC, Data and Network Science
Step 1: Synthetic populations

- **Who, where, what, when:** *People*
  - Individuals
  - Household structure
  - Statistically identical to U.S. Census
  - Assigned to Home and Activity Locations

Beckman et al. Transportation Science, NISS technical reports, Barrett et al. TRANSIMS technical reports
Step 2: Urban dynamic social contact network

- Demographically match schedules
- Assign appropriate locations by activity and distance
- Determine duration of interaction
- Generate social network

People Vertex:
- age
- household size
- gender
- income ..

Location Vertex:
- \((x,y,z)\)
- land use.
- Business type

Edge labels
- activity type: shop, work, school
- (start time 1, end time 1)
- (start time 2, end time 2)
Social Contact Networks are not easy to shatter

Vaccinating (quarantining) high-degree people

Closing down high-degree locations

Size of the largest component after removing people of largest degree

Size of the largest connected component after closing down locations of degrees: a value

Fraction of nodes in largest component

Fraction of nodes removed in decreasing order of degree
Realistic Social Contact network differ from “simple” random networks

Portland Network:
- Cliques within same age group (0-19).
- Such new measures can only be produced by Simdemics like models
Research Questions: Relational Networks

- **Question 1:** Efficient algorithms for pattern matching in relational networks
  - Extracting relational structures as the network evolves due to interventions: *sparsification* (rather than densification) of the network

- **Question 2:** Models and algorithms for embedding detailed sub-networks
  - Often times one has detailed information about certain sub-networks, e.g. critical workers, --- public health versus critical worker health
  - E.g. embedding more detailed networks of schools and universities

- **Question 3:** Integrating new information to improve the quality of synthetic networks
  - Cell-phone based movement data
Step 3. Within Host Disease Models

Disease can be spread from one person to another.

The probability of transmission can depend on:
- type of disease
- duration and type of contact
- person’s characteristics
  - age, health state, etc.

Within host model: Probabilistic timed transition systems (PTTS)
## Three Simulations Disease Transmission

<table>
<thead>
<tr>
<th>Distinguishing Features</th>
<th>EpiSims (Nature’04)</th>
<th>EpiSimdemics (SC’09)</th>
<th>FastDiffuse</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Solution Method</strong></td>
<td>Discrete Event Simulation</td>
<td>Interaction-Based Simulation</td>
<td>Combinatorial + discrete time</td>
</tr>
<tr>
<td><strong>Performance 180 days 9M hosts &amp; 40 proc.</strong></td>
<td>~40 hours</td>
<td>2 hours</td>
<td>Few minutes</td>
</tr>
<tr>
<td><strong>Co-evolving Social Network</strong></td>
<td>Can work</td>
<td>Works Well</td>
<td>Works only with restricted form</td>
</tr>
<tr>
<td><strong>Disease transmission model</strong></td>
<td>Edge as well as vertex based</td>
<td>Edge as well as vertex based (e.g. threshold functions)</td>
<td>Edge based, independence of infecting events</td>
</tr>
</tbody>
</table>
Step 4: Model for disease transmission
EpiSimdemics Algorithm

- Generate the population
- Set initial infections
- Based on activities move the people to the locations
- Compute interactions among the people at the locations
- Some exposed people may become infected
- After their activities, the people are moved back to their home PE
- Update state of person at his home PE
Percolation Based Viewpoint of Disease Transmission

\[ ps(R) = p(1,3)p(3,2)(1 - p(1,2))^2(1 - p(2,4))^2(1 - p(3,4))^2 \]

- \( t=0 \)
- \( t=1 \): \( p(1,3)(1-p(1,2)) \)
- \( t=2 \): \((1-p(1,2))p(3,2)(1-p(3,4))\)
- \( t=3 \): \((1-p(2,4))(1-p(3,4))\)
- \( t=4 \): \((1-p(2,4))\)

- Susceptible
- Infectious during 1\(^{st}\) and 2\(^{nd}\) step
- Recovered

\[ \rightarrow \] transmission
\[ \rightarrow \] No transmission
Algorithm FastDiffuse

Graph $G$ and disease model

Choose random length $l(e)$ of each edge $e$ of $G$

Random length for edge $e=(u,v)$:

$$
\ell(e) = \begin{cases} 
  i \in \{1, \ldots, S(u)\}, & \text{with probability } (1 - p(e))^{i-1} p(e); \\
  \infty, & \text{with probability } (1 - p(e))^{S(u)}. 
\end{cases}
$$

One random run of simulation

Compute distances from $s$ in w.r.t. length function $l()$

$$
V_t = \{v : \text{dist}_{\ell}(s, v) = t\}
$$

$$
I = \{e = (u, v) : \ell(e) = \text{dist}_{\ell}(s, v) - \text{dist}_{\ell}(s, u)\}
$$

$p(e) = \text{transmission prob. on edge } e$

$S(u) = \text{infectious duration of node } u$
Correctness and Efficiency: FastDiffuse

\[ \text{prob} = p^6(1-p)^9 \]

\[ \text{prob} = p^7(1-p)^7 \]

\[ I = \{ (1,3), (3,2) \} \]

\[ \text{Same output} \]

\[ R = ( (V_0, V_1, V_2, U), I ) \]

**Theorem:**
- FastDiffuse simulates a DES and runs in linear time
Visualizing the spatio-temporal diffusion

Spatial and Temporal details on spread of disease at this scale and fidelity
Research Questions: Faster Algorithms & Generalized Diffusions

**Question 3:** Scale and Speed

- Peta-scale systems & global populations
- Animal/Birds disease: e.g. avian influenza, foot and mouth

**Question 4:** Generalized diffusions

- Generalized contagion models: (e.g. Dodds and Watts)
- Other social processes (e.g. fads, ideas, viral marketing, financial cascades...)
- Similar ideas might be applicable to happiness, obesity and smoking


Network Dynamics and Simulation Science Laboratory
Step 5: Study Effects of Interventions

- Specifying a Situation (Scenario)
  - E.g. How to represent cascading failures?

- Kinds of Interventions
  - PI: Vaccines and Anti-viral, Anti-biotic
  - NPI: Social distancing, quarantining

- Specifying an Intervention
  - When, where, whom & how much

- Cost Functions
  - Human suffering averted
  - Time gained (delay of exponential growth)
  - Resource constraints

Mathematical Model: POMDP & n-way games
Behaviors and Disease dynamics can be cast as generalized reaction diffusion: Leads to coupled networks

Co-evolving dynamical systems
Step 5 Mathematical Model: POMDP & $n$-way games

- **Partially Observable Markov Decision Process (POMDP)**
  - Useful in modeling sequential decision making, e.g. policies
  - POMDP is exponentially larger than the natural problem representation (coupled dynamical system)
  - Capture the co-evolution of social network & actions

- **Mechanics**
  - Each step decision maker makes partial observations
  - Interventions based on the observation (NPI or PI: vaccinate)
    - Changes network structure or individual attributes, behaviors
  - Disease advances by one or more steps
  - Compute Cost (based on action, system state), e.g. #people infected
Effects of early detection and targeted interventions (Nature’04)
Case Study: Co-evolution of network, public policy and individual behavior
Interventions

- **Static Non-Adaptive Interventions**
  - Vaccination/AV = Static changes to social network
  - Useful as a general precautionary measure and when vaccines are very effective

- **Policy based triggered interventions**
  - School closure = One time dynamic change triggered by an event

- **Modify Individual Behavior (Decisions made by individuals)**
  - Improve individual adherence = Continuous changes to network dependent on disease status
Vulnerable versus Critical Individuals

- **Vulnerability**($\nu$): $\Pr(\nu = \text{infected})$
- **Criticality**($\nu$): reduction in epidemic size when the node is vaccinated
- Vulnerability $\neq$ Criticality.
- Criticality and vulnerability of a node do not depend only on demographics or contacts (degree)
- Vaccinating vulnerable nodes is effective

*Blue nodes are highly critical but not very vulnerable*
Triggered Policy + Individual behavior: curtailing non-essential activities

- A triggering event causes a proportion of the population to eliminate non-essential activities

<table>
<thead>
<tr>
<th>Proportion Changing Behavior</th>
<th>Final Attack Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>66%</td>
</tr>
<tr>
<td>10%</td>
<td>61%</td>
</tr>
<tr>
<td>20%</td>
<td>55%</td>
</tr>
<tr>
<td>30%</td>
<td>50%</td>
</tr>
<tr>
<td>40%</td>
<td>44%</td>
</tr>
<tr>
<td>50%</td>
<td>38%</td>
</tr>
<tr>
<td>60%</td>
<td>33%</td>
</tr>
<tr>
<td>70%</td>
<td>27%</td>
</tr>
<tr>
<td>80%</td>
<td>24%</td>
</tr>
<tr>
<td>90%</td>
<td>24%</td>
</tr>
<tr>
<td>100%</td>
<td>24%</td>
</tr>
</tbody>
</table>
DIDACTIC: HPC Services Based Epidemiological Planning Environment

Analyst can focus on delivering results rather than becoming a computing expert
Simple User Interface to Set up Experiments

Highly resolved parameters

- Population
- Disease
- Initial conditions
- Interventions
  - Type
  - Efficacy
  - Compliance
  - Timing
  - Subpopulation
Wireless Epidemics
Wireless Epidemics

- Ubiquity of smart digital devices have amplified opportunities for malware attacks
- 20.7 million devices shipped in US in 2007
  - China, India together have approx 1 Billion devices
- First generation worms like Cabir, Mabir, CommWarrior
  - No significant damage, but attacks are predicted to increase
- Wireless Epidemic can bring human epidemics closer to Internet epidemics
  - Spatial diffusion
Models

- Classical Approaches
  - Exact simulation studies (NS-2)
    - Use detailed worm models, small sized networks, simple mobility models (e.g. Random Waypoint)
    - 24-48 hours for 50 node simulation!
  - Classical ODE studies

- EpiNet: Scalable Networked Models
  - Detailed models for mobility, device assignment and activity
    - Realistic coupled social + wireless networks!
  - Provably approximate representations of Blue tooth worms
  - HPC based methods for scalability

From “The Wireless Epidemic,”
Effect of Interventions

Solution: Better late than never & a stitch in time saves nine!
Conclusions and Summary

- **Simdemics: Network-based Computational Epidemiology**
  - Highly resolved, captures complex social and epidemic interactions
  - Case Studies demonstrate practical usefulness and guide R&D
  - Extensions to wireless epidemics
  - HPC, Network and Data Science, Algorithms all play crucial role

- **Suggests new directions for future research**
  - Information extraction from relational networks
  - Understanding other social diffusion processes (e.g. economic contagion)
Related References


3. C. Barrett, S. Eubank, V. Anil Kumar, M. Marathe. Understanding Large Scale Social and Infrastructure Networks: A Simulation Based Approach, SIAM news, Mar 04.


Related References


Related References


Additional References

Additional References