Decision and Policy Informatics for Large Co-Evolving Socio-Technical Networks

Madhav Marathe
(with Chris Barrett and Stephen Eubank)

Dept. of Computer Science &
Network Dynamics and Simulation Science Laboratory
Virginia Bioinformatics Institute
Virginia Tech
NDSSL TR-09-040
These slides are a slightly extended version of the invited presentation given on April 28 2009 at Linkoping, Sweden. 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.


CRIS 2009: Fourth International Conference on Critical Infrastructures

Venue & Date: Linkoping, Sweden, April 28 2009.
Acknowledgements: Virginia Tech: Network Dynamics & Simulation Science Laboratory, VBI &

Work funded in part by NIGMS, NIH MIDAS program, CDC, Center of Excellence in Medical Informatics, DTRA CNIMS, NSF, NeTs, NECO, OCI program, and VT Foundation.
Policy Informatics for real world co-evolving networks

- Increasing interdependence and coupling between socio-technical systems caused by
  - technology dependence
  - escalating use of resources by a growing population

- Policymaking in the event of a cascading crisis or sustainable plans
  - E.g. Swine Flu outbreak, Financial meltdown and its side effects, Future energy systems

- Advances in computing and information technology → qualitative change for supporting public policy
  - New methods for collecting and integrating data
  - Service oriented methods
  - High-performance computing systems
Technical difficulties

- Networks need 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 and networks

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

Unique challenge for researchers in HPC, Data and Network Science
Interactors & Graph Dynamical Systems

- A finite undirected graph $Y$
- A sequence of local maps
- An ordering of the vertex set of $Y$

$$[F_Y, \pi] = \prod F_{\pi(i)}$$
Pervasive Cyber-environment

- Change ways in which HPC-based models and analytical tools are delivered to analysts
  - Make HPC resources seamless, invisible for routine analytical efforts
  - Organize HPC resources as an evolving commodity

Analyst can focus on delivering results rather than becoming a computing expert
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
Interaction Based Models

- Interaction based HPC-models
- A collection of interoperable simulations of societal infrastructures
- Coupled with individual-based social networks
- Individual based realistic behavioral models
  - Who, What, Where, When, and How
- Unprecedented Scale and Resolution: 300 million individuals, 100 million locations, temporal scale of minutes and spatial scale of few meters

Represent individual infrastructures, their inter-dependencies and their interactions with individuals
Cyber-infrastructure Coordination

Cyber-Infrastructure Supporting Complex Systems Research

Simfrasstructure: Service-oriented HPC Architecture

- **EpiSimdemics**
  - Interaction based model
  - Dynamic social network
  - Dynamic interventions
  - Slow compared to EpiFast

- **EpiFast**
  - Percolation based model
  - Static interventions
  - Static social network
  - Fast compared to EpiSimdemics

Data Grid
- Data Warehouse
- Data Grid
- Data Augmentation
- Data Retrieval
- Data Developer

Data Management
- Digital Library
- Generic GUIs provide user interface
- Pass to Broker

Service Broker
- Receives requests from all sources and organizes them and negotiates with other “brokers” to fulfill the requests as efficiently as possible

Model Broker

Compute Grid
- DoD HPC
- Compute Brokers

Network Dynamics & Simulation Science Laboratory
## Synthetic Information

<table>
<thead>
<tr>
<th></th>
<th>Individual</th>
<th>Aggregate</th>
<th>Networked</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relational or Nominative</strong></td>
<td>Age, income, device type</td>
<td>Number of infected individuals</td>
<td>Share a car, meet for 2 hours</td>
</tr>
<tr>
<td><strong>Declarative</strong></td>
<td>Individual activities, e.g. shop</td>
<td>Close school and work</td>
<td>Shared activities, e.g. talk on phone</td>
</tr>
<tr>
<td><strong>Procedural</strong></td>
<td>Self isolation</td>
<td>Guard sequestration</td>
<td>Driving, packet transfer protocol</td>
</tr>
</tbody>
</table>

With a given context, composing procedural and networked information allows integration of data and construction new relationships that are not measured or cannot be observed.
Decision Informatics

- Methods to translate high-level questions to appropriate workflows
  - Semantic Grid and beyond
- Data and Visual Analytics
  - Consider framing effects, confirmation bias and model accessibility
Computational Epidemiology & Public Health
Preparing for pandemics

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

**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
Simdemics: High resolution network-based modeling

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

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)
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.

**Step 3. Within Host Disease Models**

Within host model: Probabilistic timed transition systems (PTTS)
## Step 4: Fast Simulations for Disease Spread

<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</strong></td>
<td>~40 hours</td>
<td>2 hours</td>
<td>Few minutes</td>
</tr>
<tr>
<td><strong>180 days 9M hosts &amp; 40 proc.</strong></td>
<td></td>
<td></td>
<td></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>
Visualizing the spatio-temporal diffusion

Spatial and Temporal details on spread of disease at this scale and fidelity
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
Simple Co-evolution in Epi Public Health

**Policy Problem**: Is there an optimum AV allocation strategy between the market-based availability and public distribution via hospitals that minimizes the attack rate and recovers the gov’t cost of AV through markets?

**Question**: How do disease prevalence and AV demand co-evolve?
Epi-Market Study Scenario

- Actual small economically challenged region
- Total AV supply is 15k: allocated between hospitals and market
  - Hospitals: *give* to symptomatic people
  - Market: *sold* to households according to demand
- Household demand: 
  \[ D_{t,h} = \frac{B_{t,h}}{P_t} (1 - e^{-\beta x_t}) \]
  - exponential in disease prevalence \((x_t)\)
  - increase in household budget \((B_{t,h})\); decrease in price \((P_t)\)
  - price is linear in remaining supply
  - \(\beta\) reflects risk aversion or prevalence elastic demand to AV.
- Different allocations strategies, 25 replicates.
Experiment Results

- Attack rate reaches minimum when hospital allocation of AV is 6K.
- Application of hospital allocation is upper bounded by attack rate. No need to allocate more to hospitals.
- Extra AV can go to the market. Revenue recovers cost.
Experiment Results: 2K vs 6K case

• 2K-case: 2K to hospitals + 13K to markets. Attack rate 13%.
• 6K-case: 6K to hospitals + 9K to markets. Attack Rate 5%.
• AV market supply: gone on day 95 in both cases. From day 80~95, 2K-case has higher prevalence, lower price - higher demand.
• After day 95, AV still available at hospitals in 6K-case; no more AV in 2K-case - epidemic out of control.
Differential Benefits for Demographic Classes
Simple User Interface for Experiments

Highly resolved parameters

- Population
- Disease
- Initial conditions
- Interventions
  - Type
  - Efficacy
  - Compliance
  - Timing
  - Subpopulation
Visual Analytics Support
SIGMA: High Resolution Modeling Environment for Networked Markets

• An interdependency aware, end-to-end modeling environment to study network constrained commodity markets (e.g. Electricity, Gas).

• Novel Features:
  – Spatially disaggregated demand modeling
    • Consumer demand is price elastic, time, activity and demographics dependent – our modeling framework captures this
    • Models are resolved at an individual consumer level, yet can scale to large regions e.g. Chicago.
  – Highly scalable
  – Integration consumers, transmission and distribution network and producers. The first two are rarely put together into market design models. Market mechanisms depend on all 3 components
  – Integrated with other infrastructures such as communication, transportation and health.
E.g. Locational Market Power


- We introduce a flow based, quantifiable definition of market power.
- We analyze topological cause of market power for the Portland electrical network.
- Strategic alliances can make the issue of locational market power much worse.
- Elastic/price responsive demand can mitigate the locational market power effect.
Locational Market Power (contd.)


- Developed analytical tool to study the cause of locational market power.
- Given the network topology, is locational power caused by the network capacity or production capacity?
- Understanding this issue helps determine where the future infrastructural investments should be made.
- Results find Portland market to be production constrained.
- Results also find significant interaction between the economic and physical components in the network constrained markets.
Renewable Energy


- Discusses the Renewable Portfolio Standard requirements for different states.
- Studies the implications of an integrated market for tradable renewable energy credits.
- Several markets (or states) are integrated into one super market which ensures environmental benefits inside the region covered by the super market territory. For example, one super market can be established for New York, New England and PJM (Pennsylvania, New Jersey and Maryland).
Electrical Grid Analysis


- Study the vulnerability of real, synthetic and random electrical networks using graph based structural analysis.
- Results suggest that all grids are more vulnerable to targeted attacks compared to random attacks.
- The electrical networks have higher vulnerability compared to other infrastructural networks and social networks.
- Targeted attack on several high capacity transmission lines is not able to affect the flow of the real network (Portland) in any significant way, implying redundancy in the transmission capacity.
Conclusions

- An integrated program to support policy and decision making vis-a-vis co-evolving socio-technical networks
- Spiral R&D program guided by practical user-defined case studies
- Future Step: Vision
  - Supporting Policy Informatics using Google++ (Active Queries that spawn models and integrate data)
  - Combine the system with surveillance information
Related References


Related References


Related References


References on Energy


