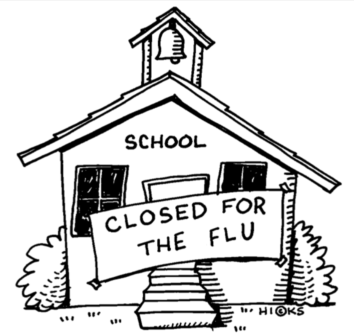
# Simulation of Contagion on Very Large Social Networks

Keith Bisset, Jae-Seung Yeom
Network Dynamics and Simulation Science Lab
Virginia Tech

# The Problem – Contagion Diffusion

- Compute the spread of contagion across large contact networks (3x10^9 - 8x10^10 nodes)
  - **Contagious Disease (Influenza**, Smallpox, Pertussis)
  - Information Flow (Rumors, Fads)
  - Social Contagion (Smoking, Obesity, Incarceration, Social Unrest)
  - Computer Malware (Viruses)
- Goal: Evaluate interventions to reduce or increase spread

# Need for Computational Epidemiology

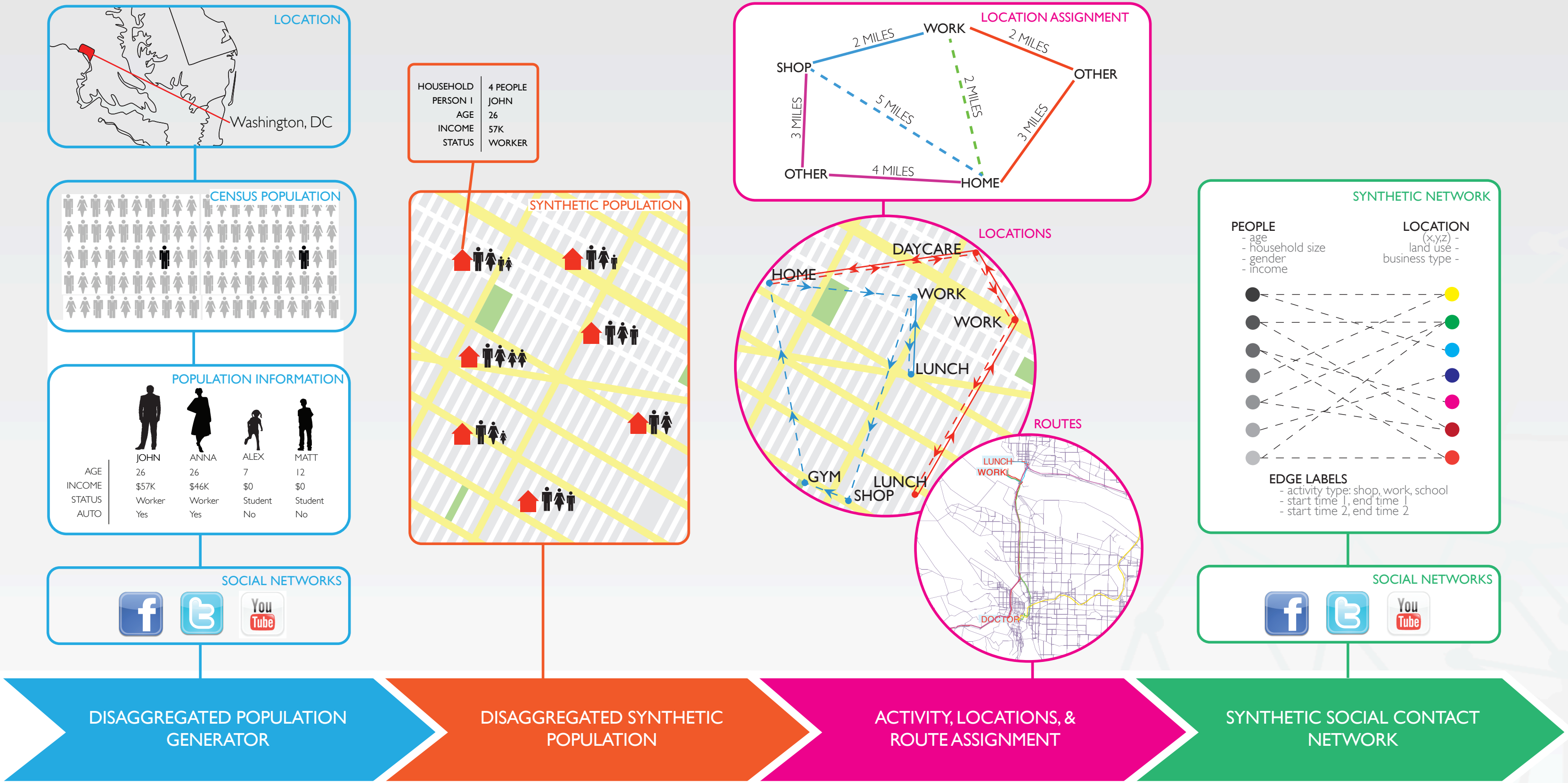



- Infectious diseases account for more than 13 million deaths a year.
- Controlled experiments often impossible in epidemiology due to ethical and practical reasons.
- Computational models help in understanding the space-time dynamics of epidemics.
- Pervasive computational environments can provide real-time access to models, data, and expert opinion to analysts as an epidemic unfolds.

# Computational Epidemiology Driven by Synthetic Information

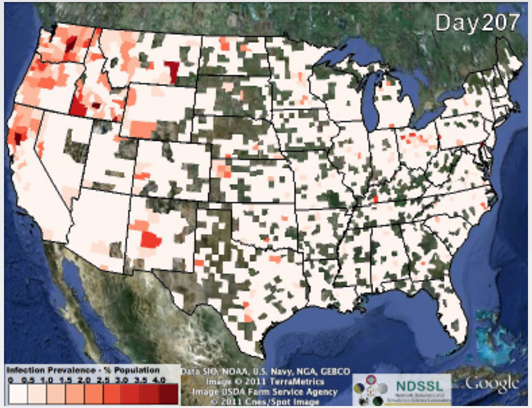
- Synthetic Information Provides
  - Interactions with other systems
  - Rapid response to emerging crises
  - Flexibility to adapt to different areas
  - Integration of disparate data sources
- Forecasting
  - Surveillance
  - Social media data
  - Blogs
  - Survey data
- Course of Action Analysis
  - Cost-benefit analysis
  - Prioritization of interventions
  - Planning Exercises
- Situational Awareness
  - Estimation of number infected
  - Estimation of disease characteristics
- Preparedness and Training

# Constructing synthetic social contact networks

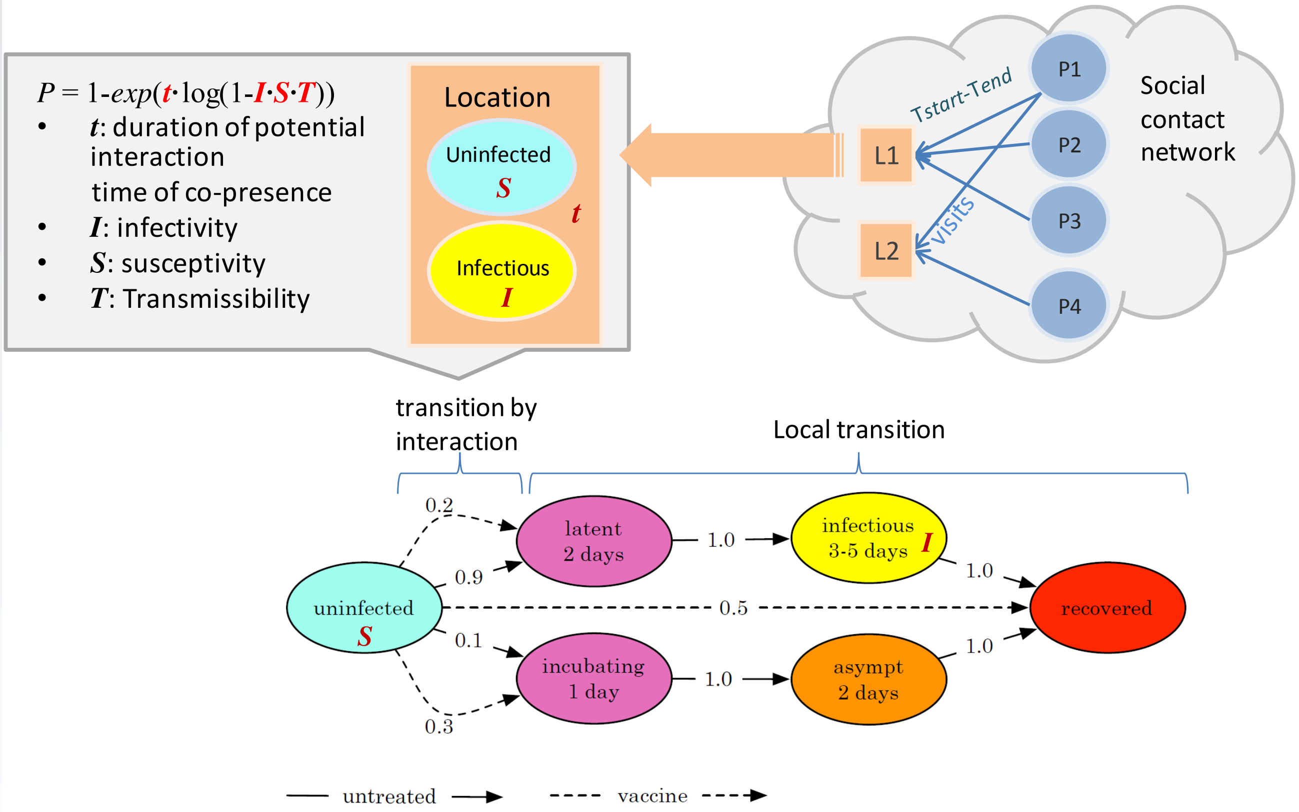



VirginiaTech
Invent the Future

# Simulating contagion over dynamic networks




- Agent-based model
- Realistic Population
- Complex Interventions
- Co-evolution between
    - Contact Network
    - Individual Behavior
    - Public Policy
    - Individual Health



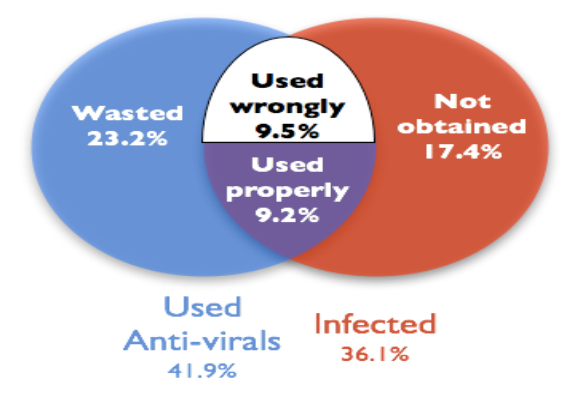
# Selected Case Studies

- Pandemic Influenza Planning – Fall 2006
  - How can we prepare for a likely influenza pandemic?
- Emergence of H1N1 Influenza – Spring - Fall 2009
  - What are the characteristics of this influenza strain and their likely impact on US populations?
- Optimizing Household Response to Epidemics – Spring 2011
  - Do behaviors within households, while harder to simulate, have a significant impact on the course of an epidemic?
- Adenovirus Pandemics Simulation and Analysis – August 2013
  - How can decision makers become familiar with the challenges and decisions they are likely to encounter during a national pandemic, mainly centered on the allocation of scarce resources?
- Emergence of Middle East Respiratory Syndrome – Now
  - We are exploring how simulation combined with a synthetic informatics analytic approach might shed light on what may be driving this outbreak.

# Antiviral Distribution Planning

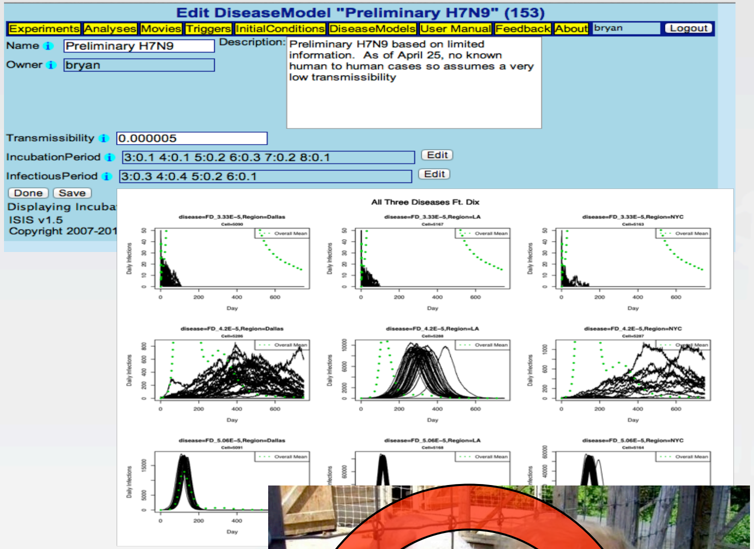
**Would retail dispensing of antivirals reduce morbidity and mortality?**

- Model individual decision-making in socio-economic context
- Incorporate the diagnostic process (false + / -) into intervention applications
- Dynamic application of antivirals based on current state of infection
- Integrating economics of health care decisions with dynamic simulation
- Effect of individual decision-making on epidemics

# Emergence of H1N1 “Swine Flu”

- Early days of 2009 H1N1 pandemic
- Transmissibility and Virulence not known
- Analysts at DTRA Reachback were running simulations and interpreting results on a 24 hour cycle to provide input to a daily briefing to HHS as the epidemic was unfolding
- Regular training of analysts played critical role in rapid response




# Open Source Data based Forecasting

**Can freely available data be integrated with statistical and simulation models to forecast the near term future?**

- Goal: *Flu Forecast* similar to current weather forcasts
- Unlike weather people can actually do something to change outcome
- Harness social media to infer population level sentiments and identify events of interest
- Processing of alternative data sources upstream from event causation
- Fusing multiple predictive data streams to dynamically optimize predictions
- Integrating time series analysis with simulation models
- Evaluating social media as an information source

# ISIS: Complete analyst workflow support

## Hypothesize

Influenza Like Illnesses — Mary | Completed — START | DUPLICATE | DELETE

Influenza Pandemics Threat — Tom | New — START | DUPLICATE | DELETE

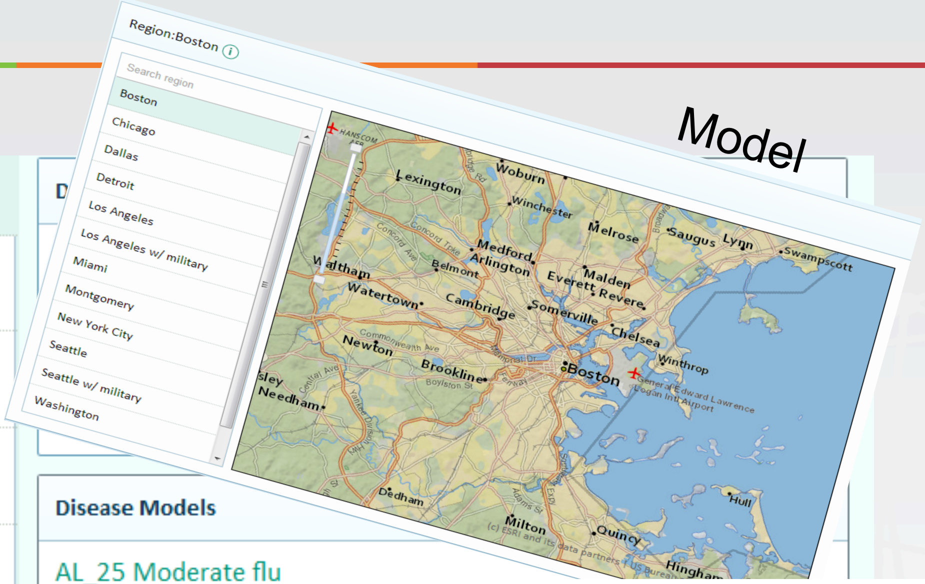
Liver Cancer — Mary | New — START | DUPLICATE | DELETE

Malaria Control Programs — Tom | New — START | DUPLICATE | DELETE

DUPLICATE | DELETE

DUPLICATE | DELETE

## Model

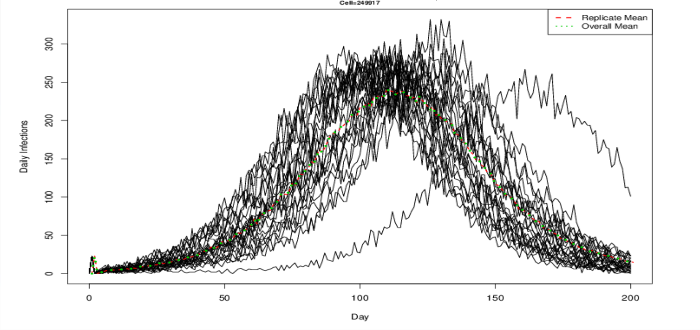



Disease Models

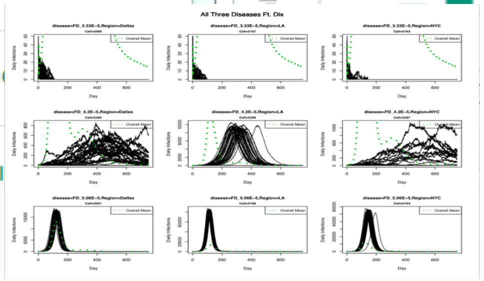
AL_25 Moderate flu

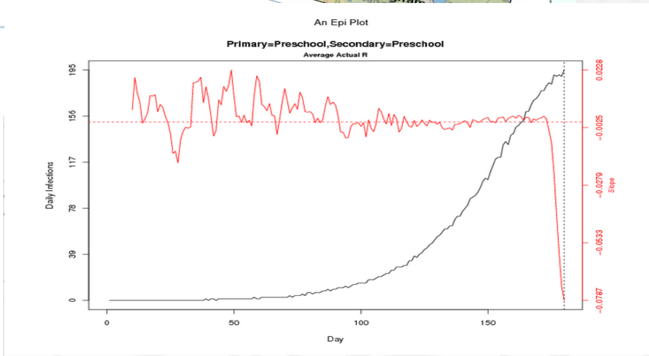
Incubation Period

## Predict, Decide




## Visualize



# Key Challenges

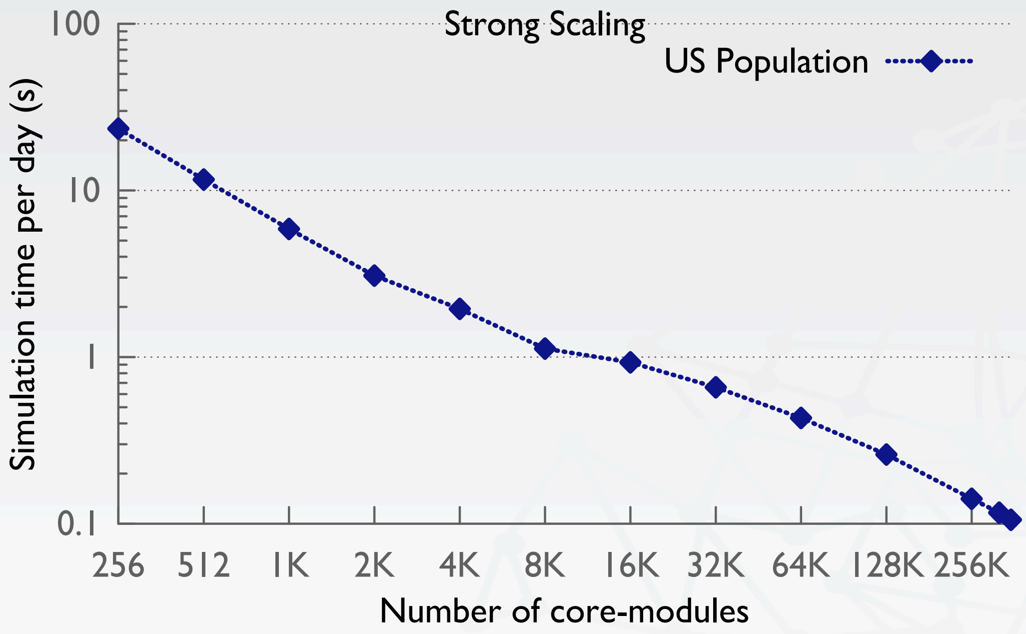
- Large, highly irregular, dynamic network
- Large number of small messages
- Large number of runs needed for study, even more for forecasts
- Atypical instruction mix (not floating point)

| Instruction Type | Portion |
|---|---|
| Load/Store (L1-D) | 58% |
| Branch | 19% (70% Conditional) |
| Floating Point | < 1% |

# EpiSimdemics Strong Scaling

## Model 280 million people on 352k cores for 120 days: 12 seconds

- System: NCSA Blue Waters
- 352k Cray XE6 core modules
- US Population: 280 million
- Simulated time: 120 days
- Walltime: 12 seconds
- Granularity: < 800 people/core




## Projected Global Population time: 6 minutes

To appear IPDPS 14

# Ongoing Work

- World Domination
- More complex behavior
- Long distance travel model
- Co-circulating diseases (e.g., Influenza and ILI)
- Complex Contagion

# Acknowledgements

- NDSSL and our collaborators
- NCSA
- Sanjay Kale and PPL
- NSF, NIH, DoD, CDC, HHS

*CNIMS Contract HDTRA1-11-D-0016-0001, NIH MIDAS Grant 5U01GM070694-11, NSF Blue Waters OCI-0832603, NSF PetaApps Grant OCI-0904844,*