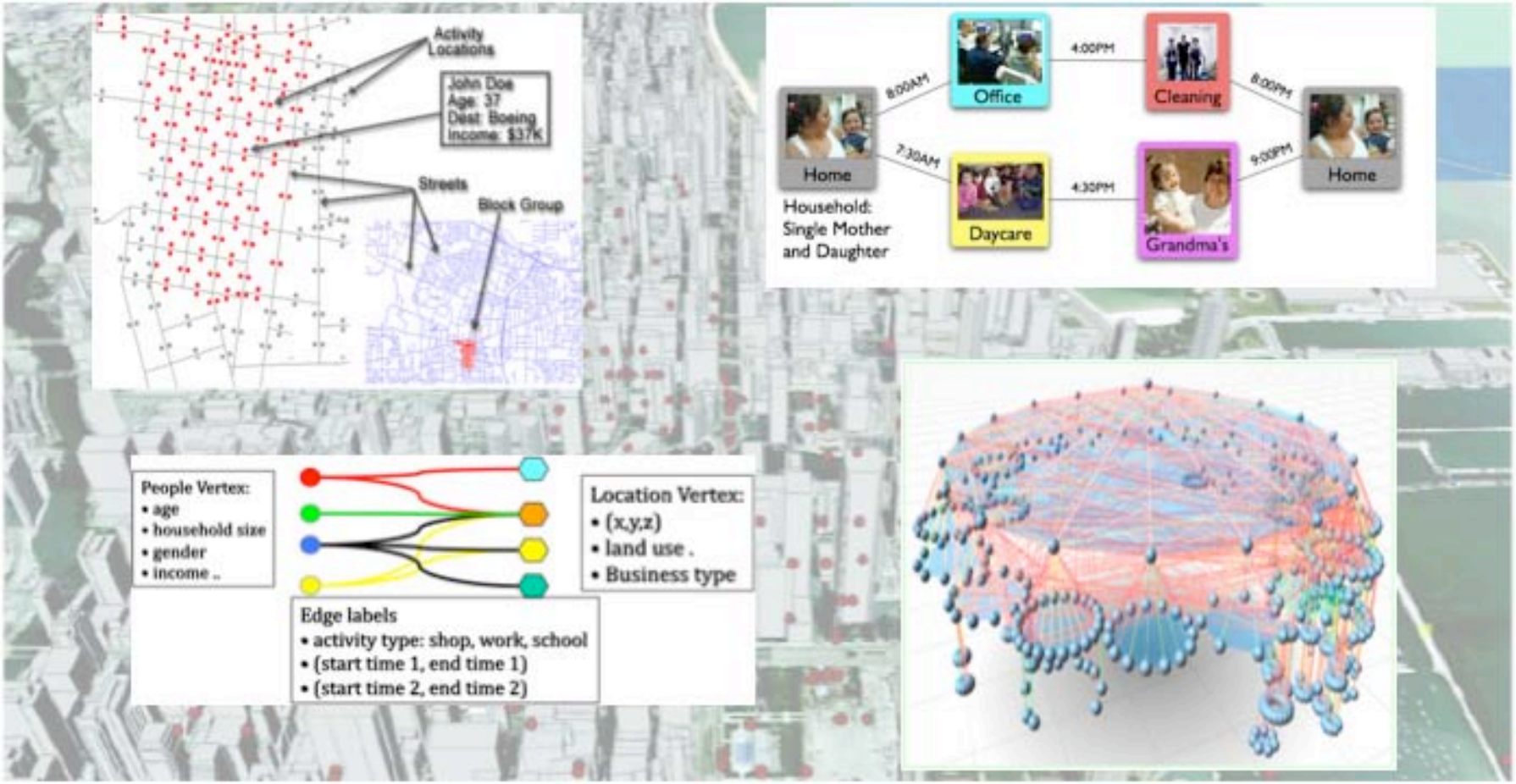
Synthetic Information and Decision Informatics for Complex Socially-Coupled Systems

Christopher L. Barrett
Network Dynamics & Simulation Science Laboratory
Virginia Bioinformatics Institute
Virginia Tech

The Problem is to Couple Social Systems with X



Data & Social Science

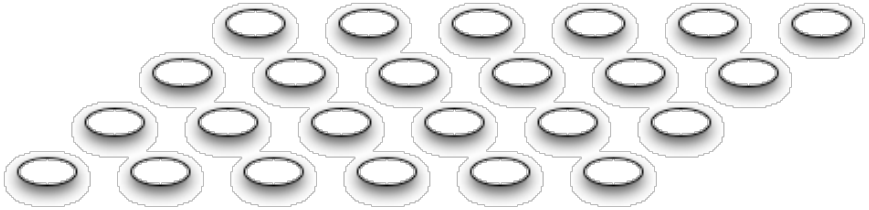
- Where do you get *that* data?
- Survey Technique and Data
  - Arguably made modern social science a science
  - "Variable rich, case poor" (-Kenneth Prewitt)
    - Control of error, reliable <u>quality</u>
    - Expensive, late, sparse
  - Never exactly what you need (-C. Barrett & everybody else)
- Administrative Data/ Public information
  - Suddenly more available electronically
  - "Case rich, variable poor"
    - Poor control of error, <u>unreliable</u> quality
    - Already paid for, current, historical, comprehensive as bureaucracy
- Privacy and other moral issues

Revolution in Social Science & Systems

- Information accessibility-based
  - Mobility: location, function, time, resolution
- Some basic dimensions
  - Mobile communication and computing
  - Service and virtual computing and data storage
  - Social computing/ computational sociology
  - Wisdom of crowds
  - Social networks and social networking
  - Inter-generational changes in how information is shared
    - conventions regarding privacy, accountability, ownership
- How can public information have survey quality?

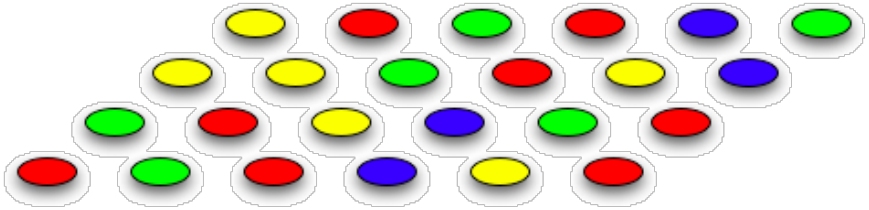
Constructing a synthetic information system: choose resolution

- Represent each interactor



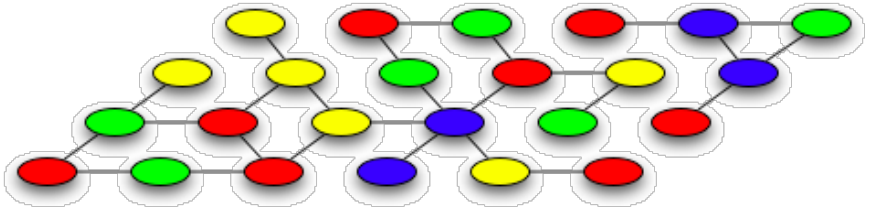
Constructing a synthetic information system: disaggregate observations

- Represent each interactor
- Endow them with individual attributes



Constructing a synthetic information system: estimate interactions

- Represent each interactor
- Endow them with individual attributes
- Generate interaction patterns

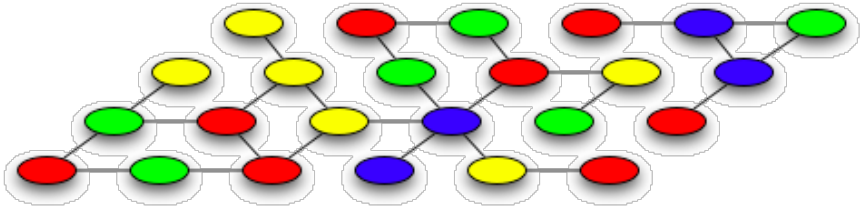


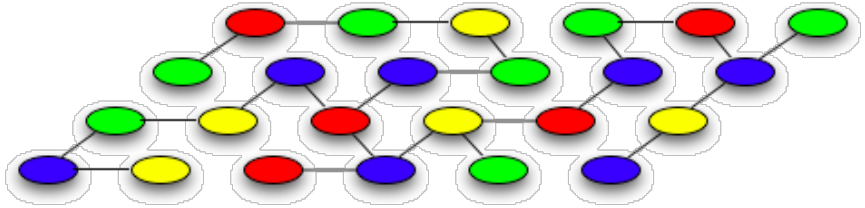
Contact / colocation network

Constructing a synthetic information system: increase fidelity

- Represent each interactor
- Endow them with attributes
- Generate interaction patterns
- Generate other dynamics

Contact / colocation network

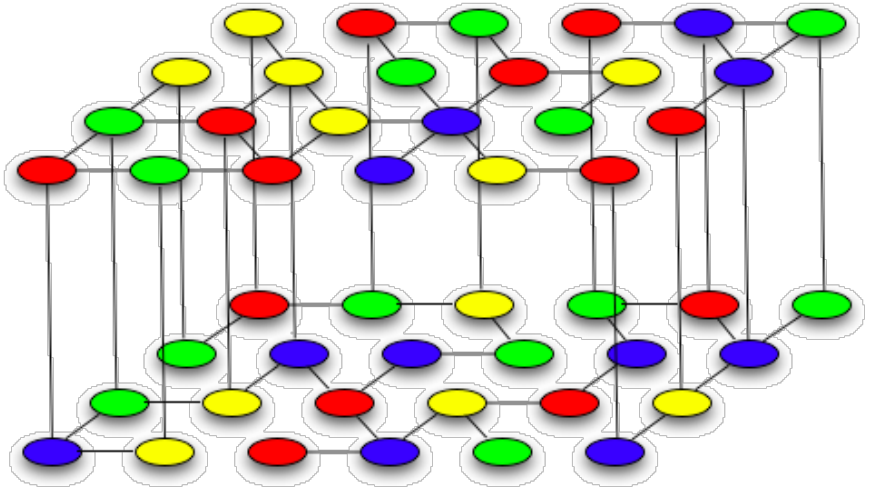




Communication network

Constructing a synthetic information system: across-level agency

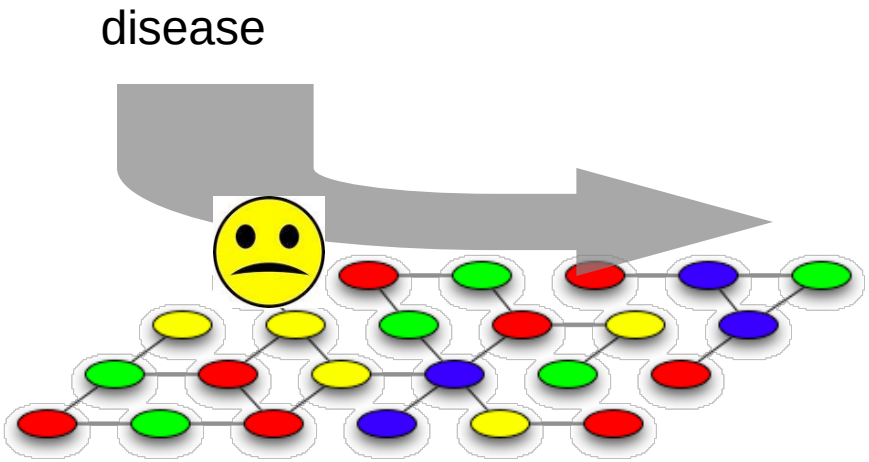
- Represent each interactor
- Endow them with attributes
- Generate interaction patterns
- Generate other dynamics
- Multi-theory composition
- Unencapsulated agency



Synthetic information systems deduce associations implied by causal hypotheses

1. Network topology affects outcome: e.g. # infected as a function of time

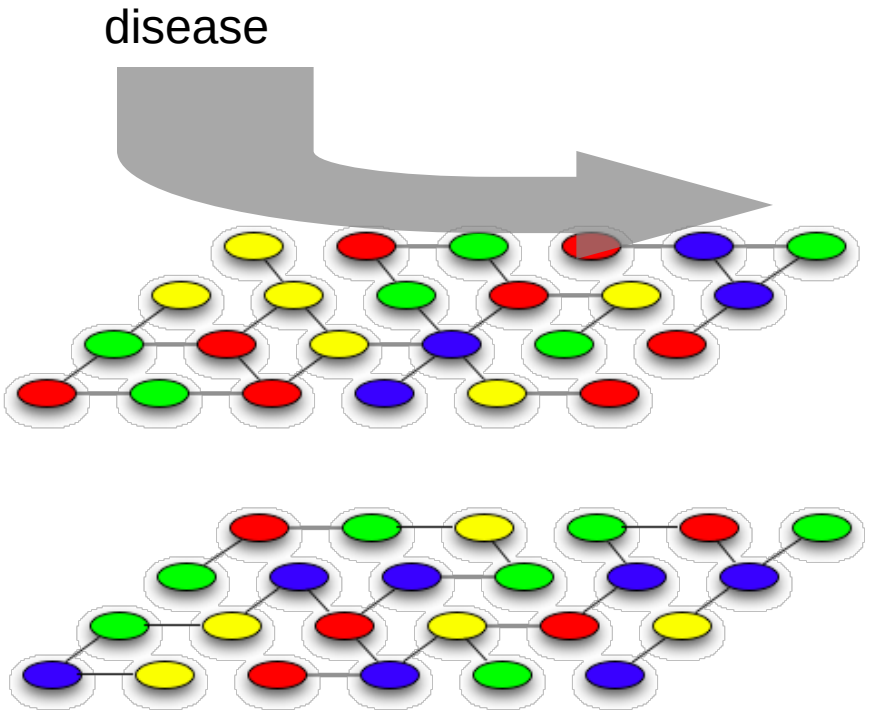
   Generates associations between attribute and probability of infection



Synthetic information systems deduce associations implied by causal hypotheses

2. Other attributes may just be "along for the ride"

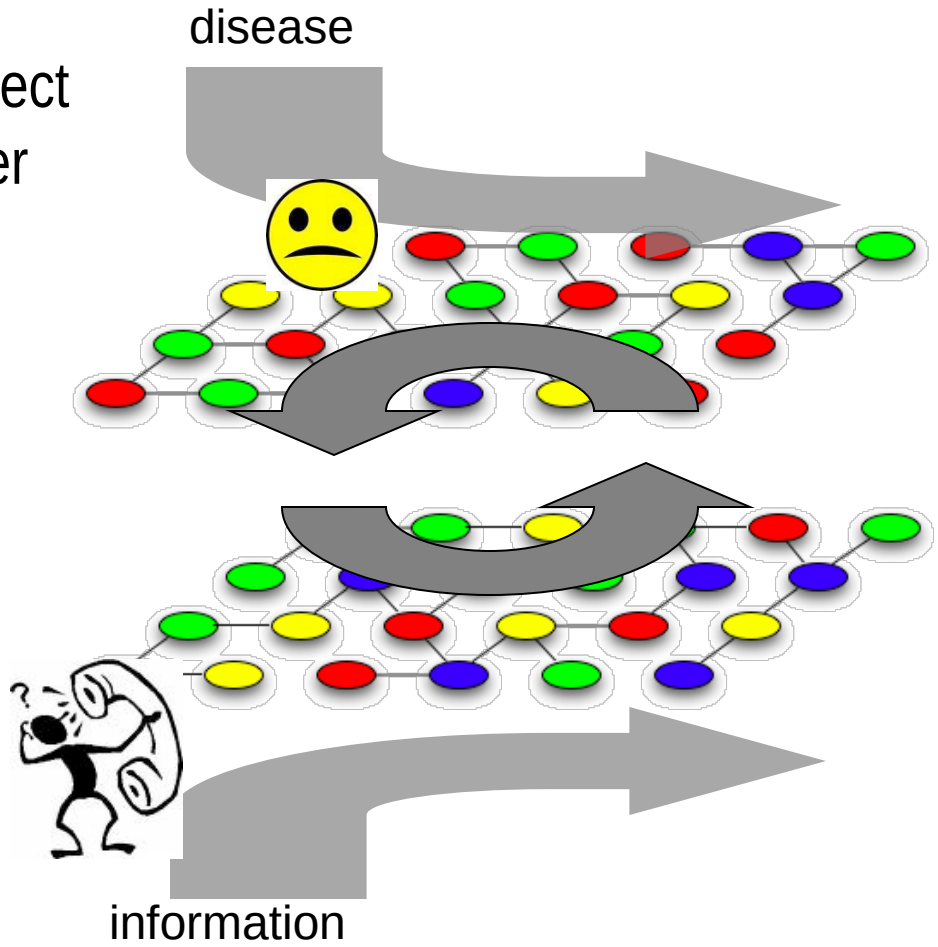
E.g. Attribute in lower network may be correlated with attributes that determine contact network



Synthetic information systems deduce associations implied by causal hypotheses

3. Interactions in one layer may affect network topology in another layer

   Associations evolve



Synthetic information systems deduce associations implied by causal hypotheses

1. Network topology and its dynamics affects outcome:

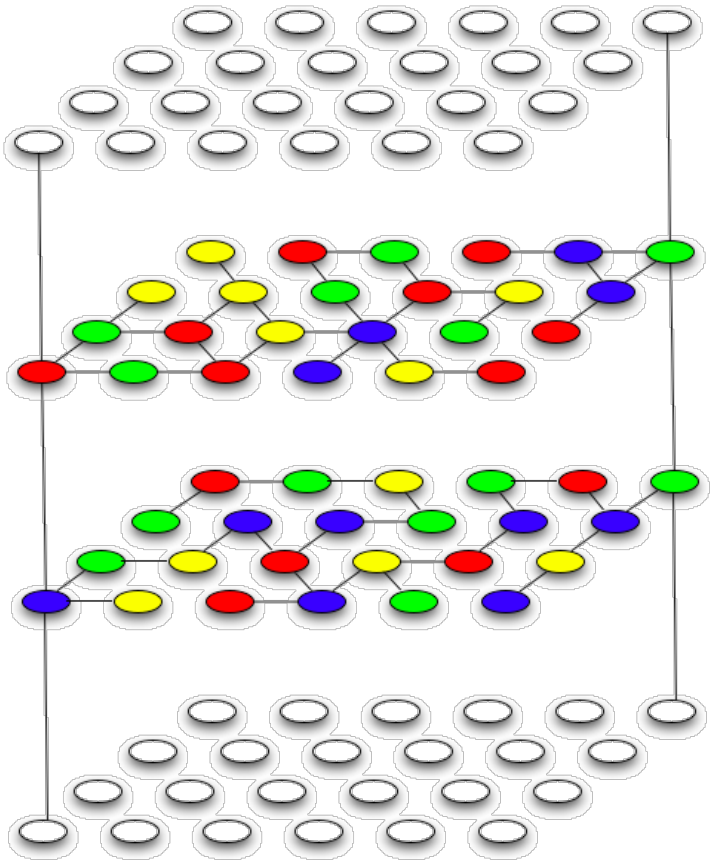
   irreducible complexity

1. Other attributes may just be "along for the ride":

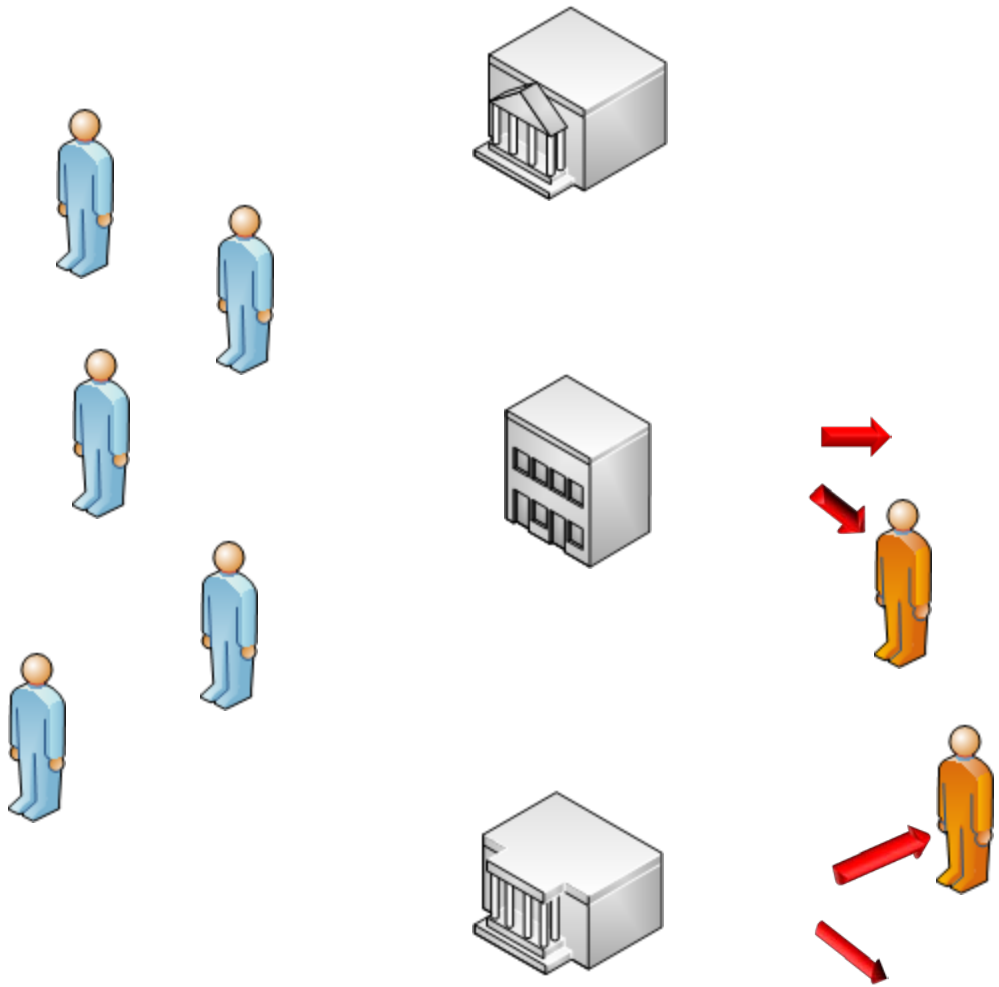
   generating confounders

2. Interactions in one layer may affect network topology in another layer:

   co-evolving networks



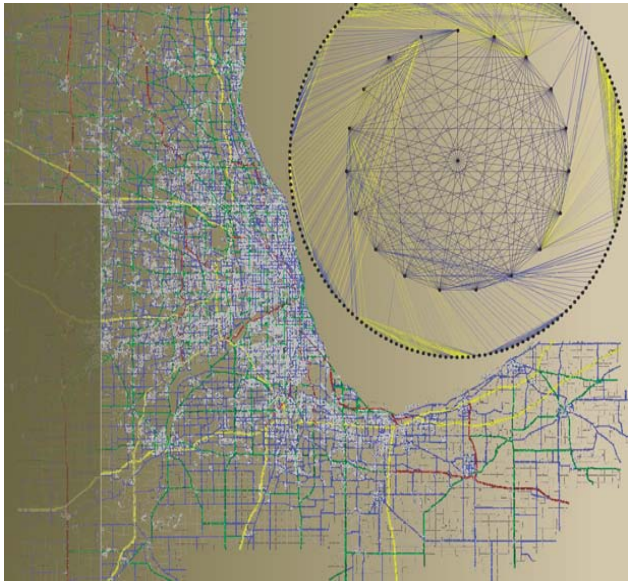
Capturing Detailed Social Interaction: Contagion example



- 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 anchor loc PE
- Update state of person at the anchor loc PE

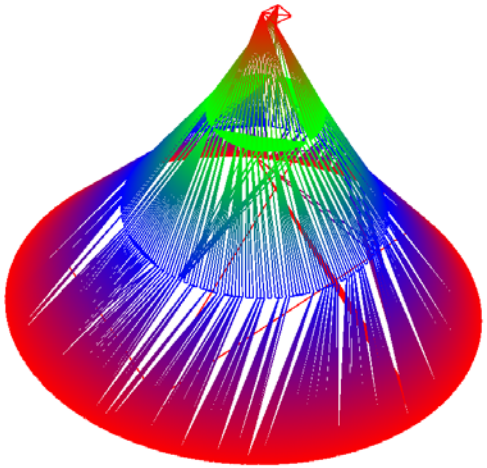
Social systems seem to generate graphs

- Usually social systems involve interacting people and locations:
  - properties, functional behaviors, constraints and competition for resources
- The interaction relationships among definable places, entities or groups,
  - Frameworks for behavior norms that internally structure the actions of (inter)actors in the social system.
  - Patterns of behavior of participants and places in relation to each other



Social *structures* seem to include graph / network labels

- Normative Structure — relational patterns in an organization among entities and conventions and that have varying societal roles/functions
- Ideal Structure — relational patterns of beliefs and views of people of varying societal roles/functions
- Interest Structure —relations between goals and objectives of entities having various social roles/functions
- Interaction Structure — structure and form of communications among entities having various social roles/functions



System Complexity, Detail & Data

- Socially-coupled systems questions are heterogeneous and asymmetric
  - Their graphs are irregular

Social System Complexity, Detail & Data

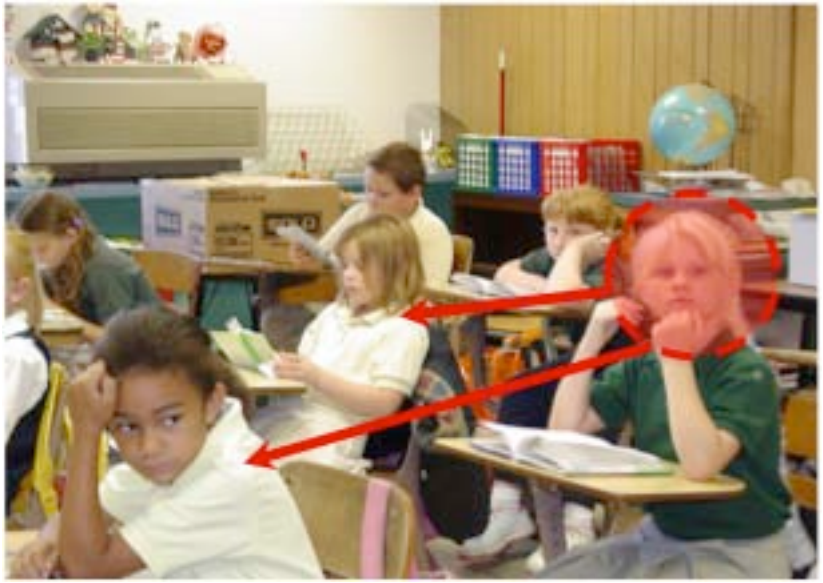
- Socially-coupled systems questions are heterogeneous and asymmetric
  - Their graphs are irregular
- Two sides of complexity
  - Irreducibility of the interaction structure and its representation
  - Irreducibility of the representation of the dynamics

System Data, Detail & Complexity

- Socially-coupled systems questions are heterogeneous and asymmetric
  - Their graphs are irregular
- Two sides of complexity
  - Irreducibility of the interaction structure and its representation
  - Irreducibility of the representation of the dynamics
- Both sides have implications with respect to HPC
  - HPC enables entirely new analytical access to big systems
  - Data sources
  - Data output
  - Compute/ recompute/store decisions
  - Algorithm design
  - Memory and storage architecture optimization
  - Compute- and data- intensive analytical methods: network-intensive computation
  - Distribution, virtual machines and usable services

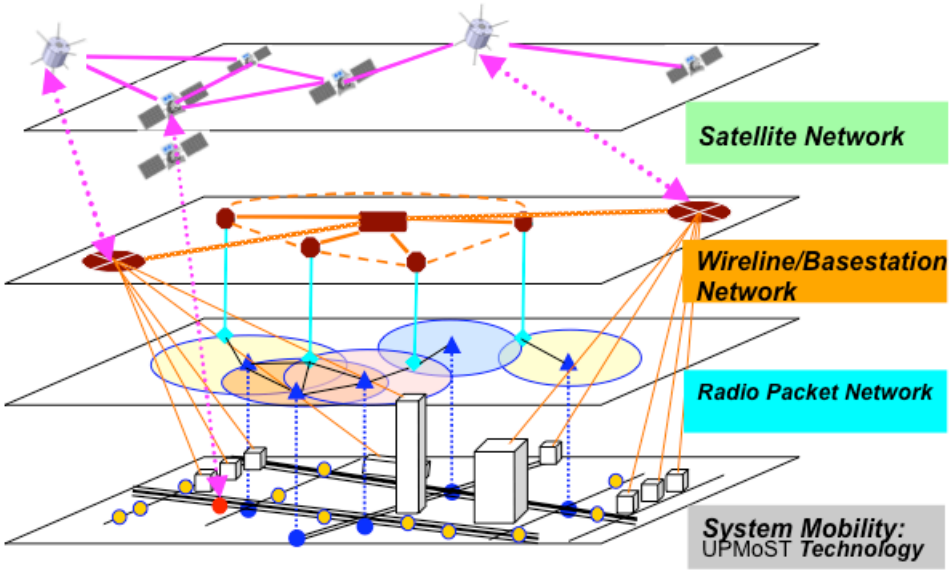
Unstructured Data and Knowledge to Synthetic Information

- The graphs/social networks are synthesized from details and make details



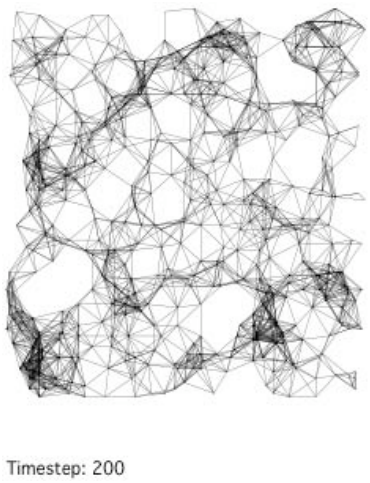
Unstructured Data and Knowledge to Synthetic Information

- Lots of things are used and synthesized
  - Nominative, including quantitative data
  - Declarative data and information
  - Procedural information
  - Relational & dynamical information



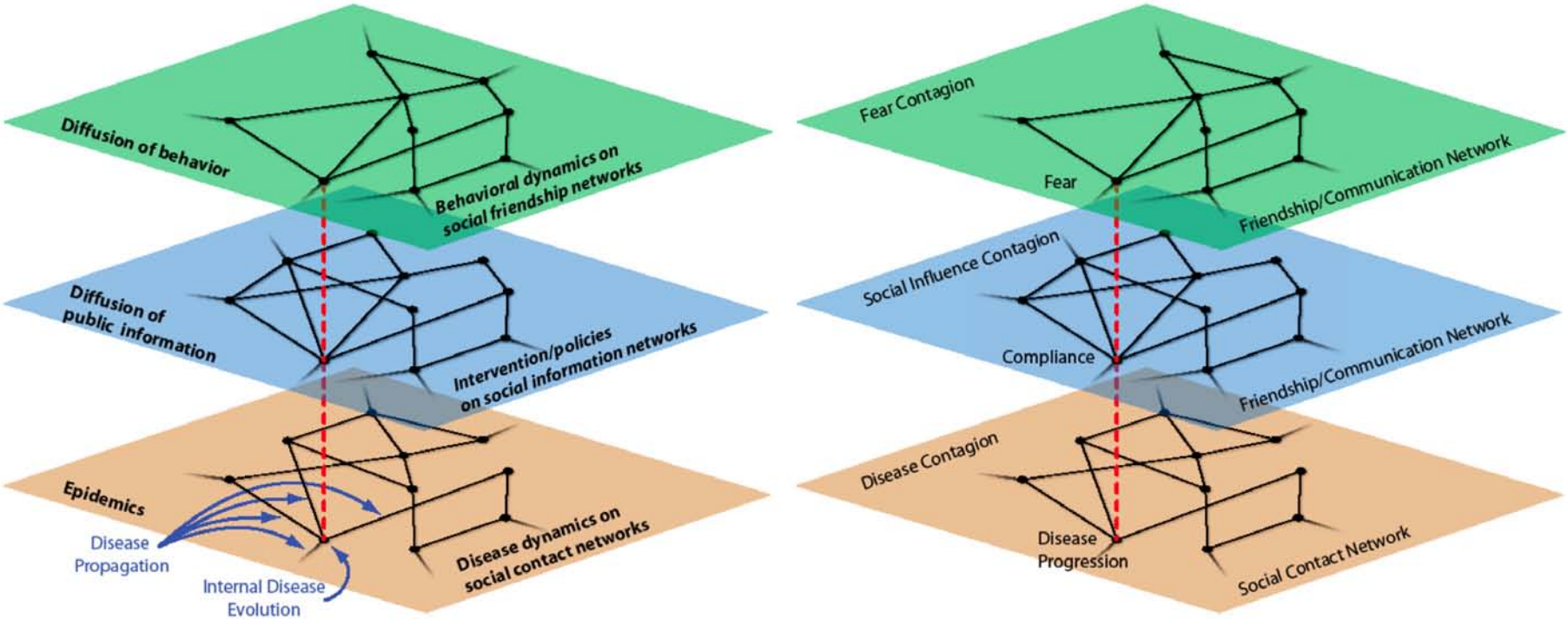
Unstructured Data and Knowledge to Synthetic Information

- What matters?
  - In depends on the problem



System Complexity, Detail & Dynamics

- Socially coupled systems co-evolve
  - That is not modeled *a-priori*, it is generated
  - Poses modeling, analytical and data issues



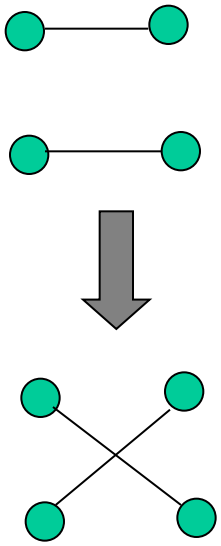
So, What Detail Matters?

- How to know what matters
  - Big issue
  - Huge data issues
  - Should replace naïve notions of (predictive) “validity” as “the question”

- For many questions in policy and real-world decision making, a lot of detail matters for socially-coupled systems
  - HPC and massive data handling is necessary

Example of a Test for Detail: Rewiring Operations

Rewiring operations

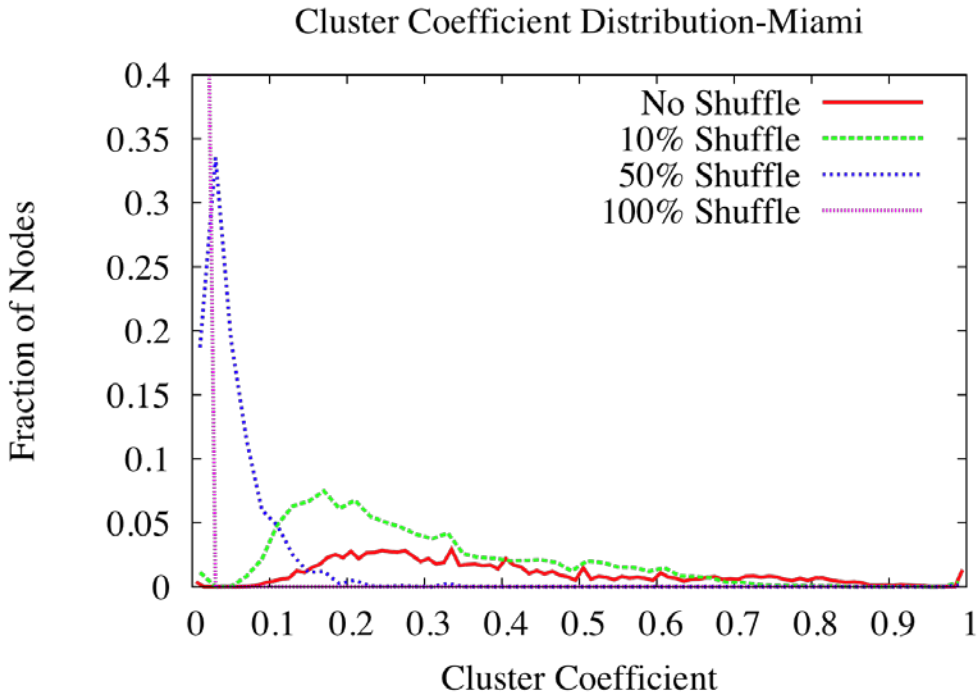


Assortativity -reserving rewiring operations

- Age assortativity: flip edges if end points have similar age
- Degree assortativity: flip edges if end points have similar degrees

Preserves degree distribution: tests influence of “scaling” alone on dynamics

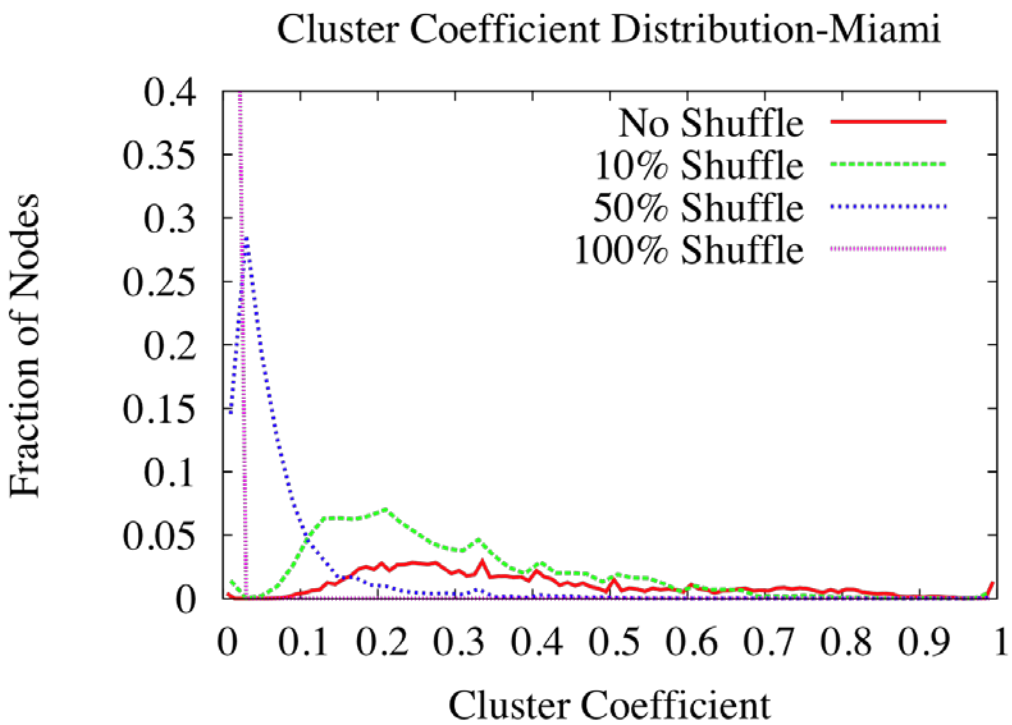
Insufficient detail: Clustering Coefficients



Effects of rewiring

- Degree distribution unchanged
- Clustering coefficient reduces significantly - similar to behavior in small world models

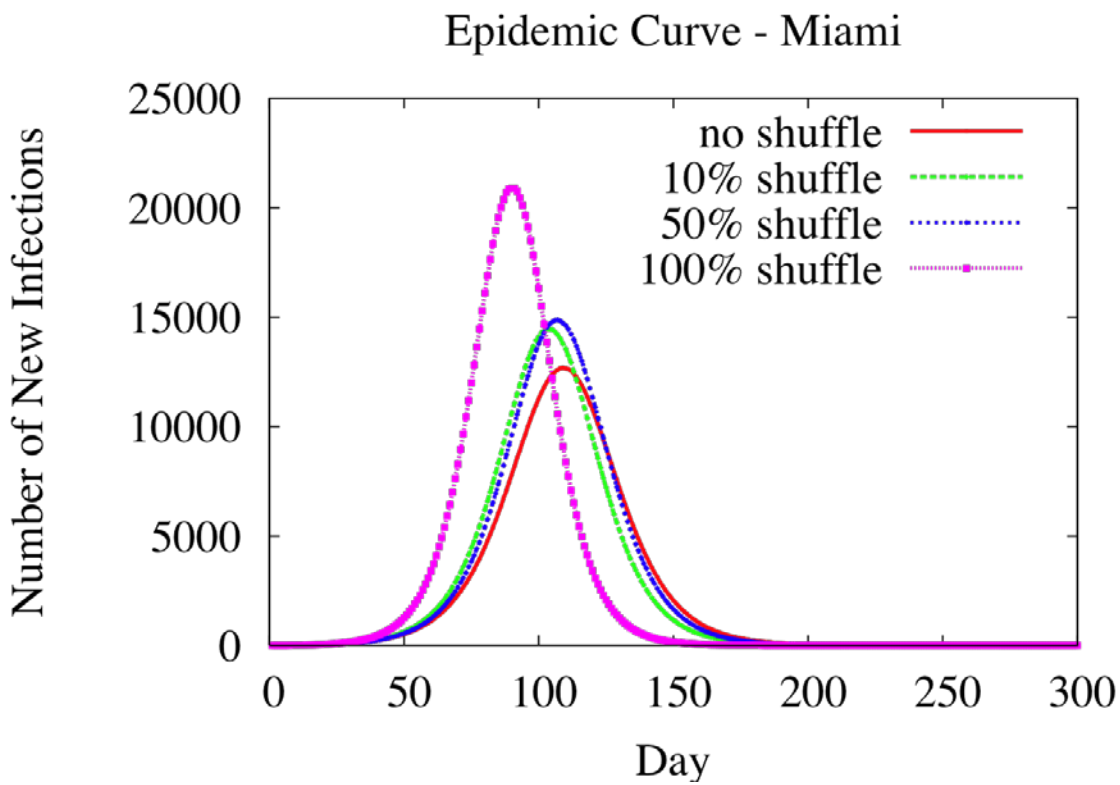
Insufficient detail: Clustering Coefficients after Age-assortative Shuffle



Effects of rewiring

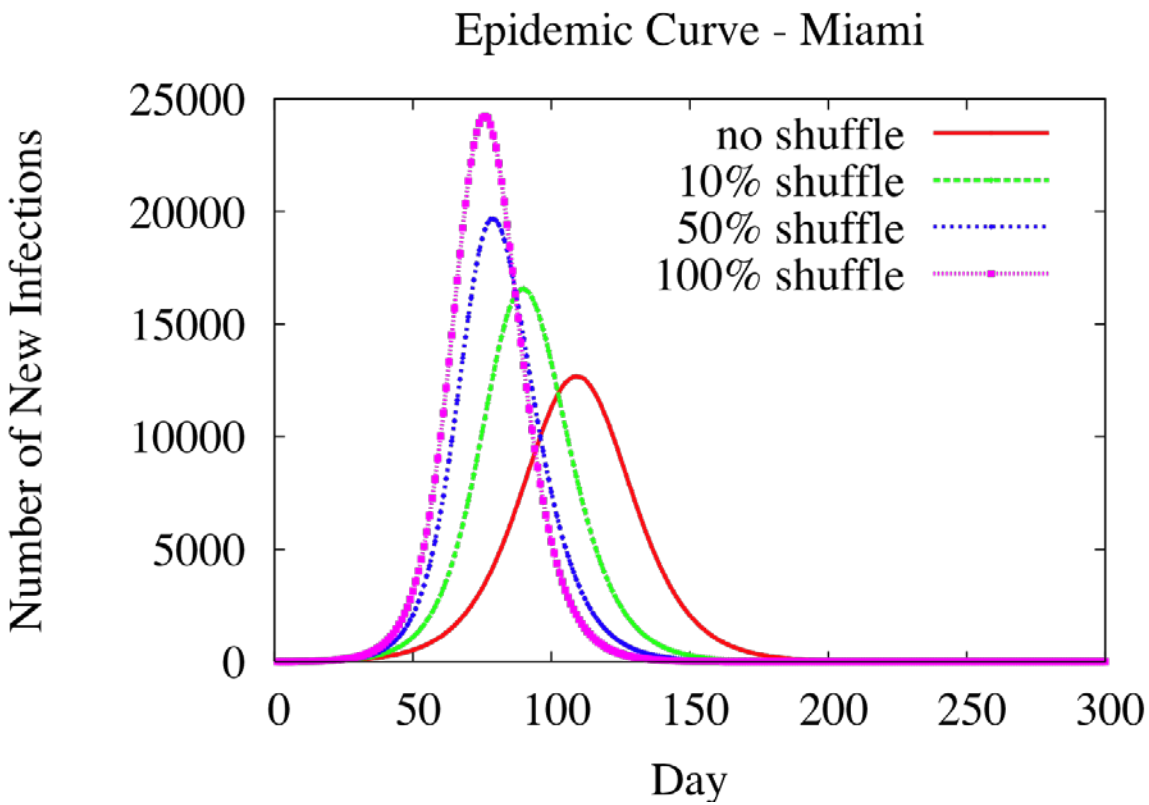
- Similar as regular shuffle

Insufficient Detail: Effect of Rewiring on Disease Dynamics



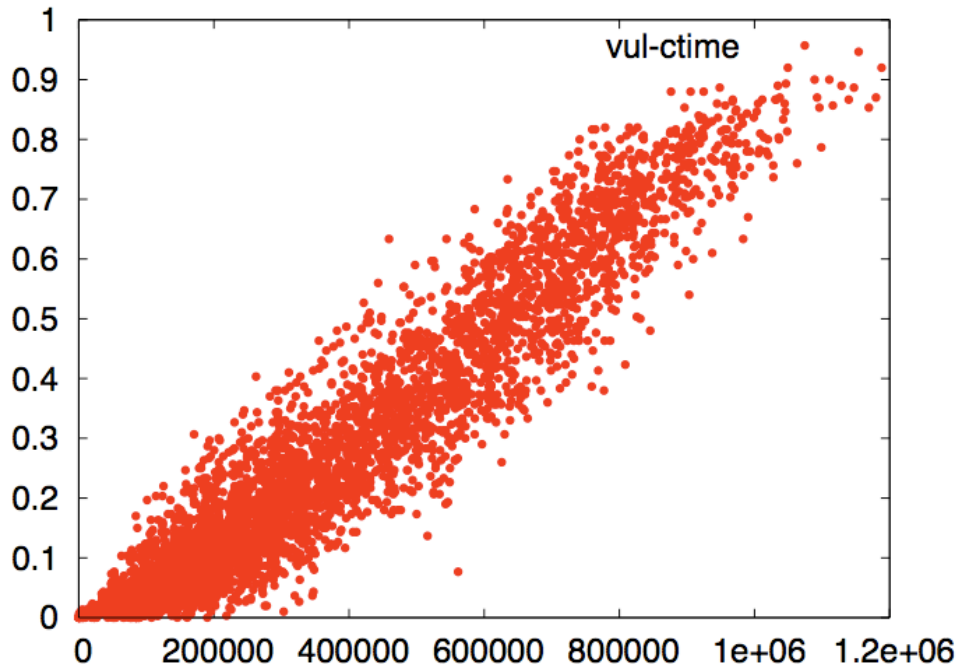
- Disease spreads quickly in the shuffled networks

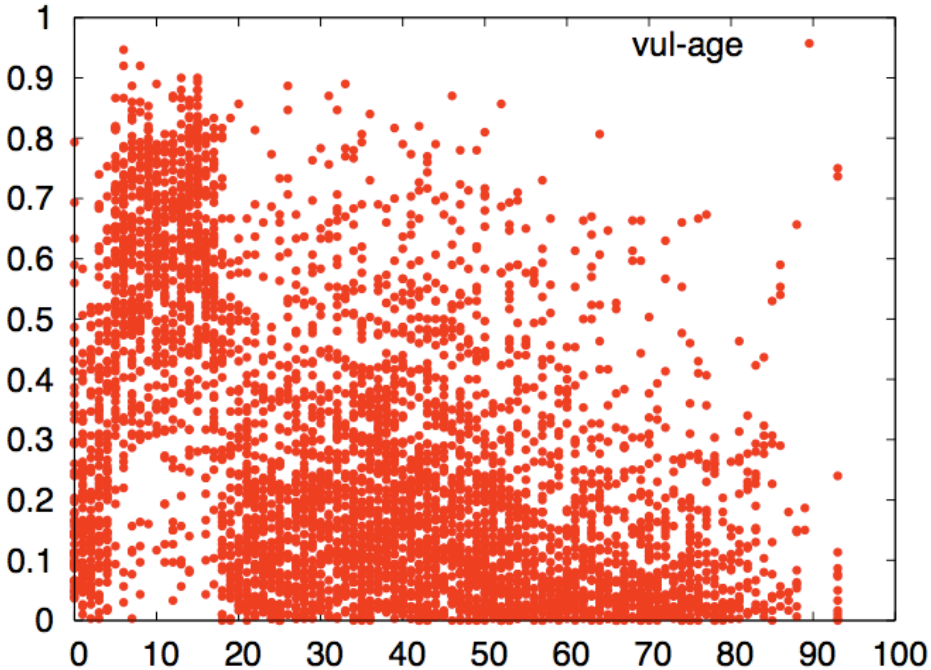
Insufficient Detail: Disease Dynamics after Age-assortative Shuffle



- In age-assortative shuffled networks, attack rates even larger than those in regular shuffled networks

Sufficiently Detailed: complex demographic & behavioral features

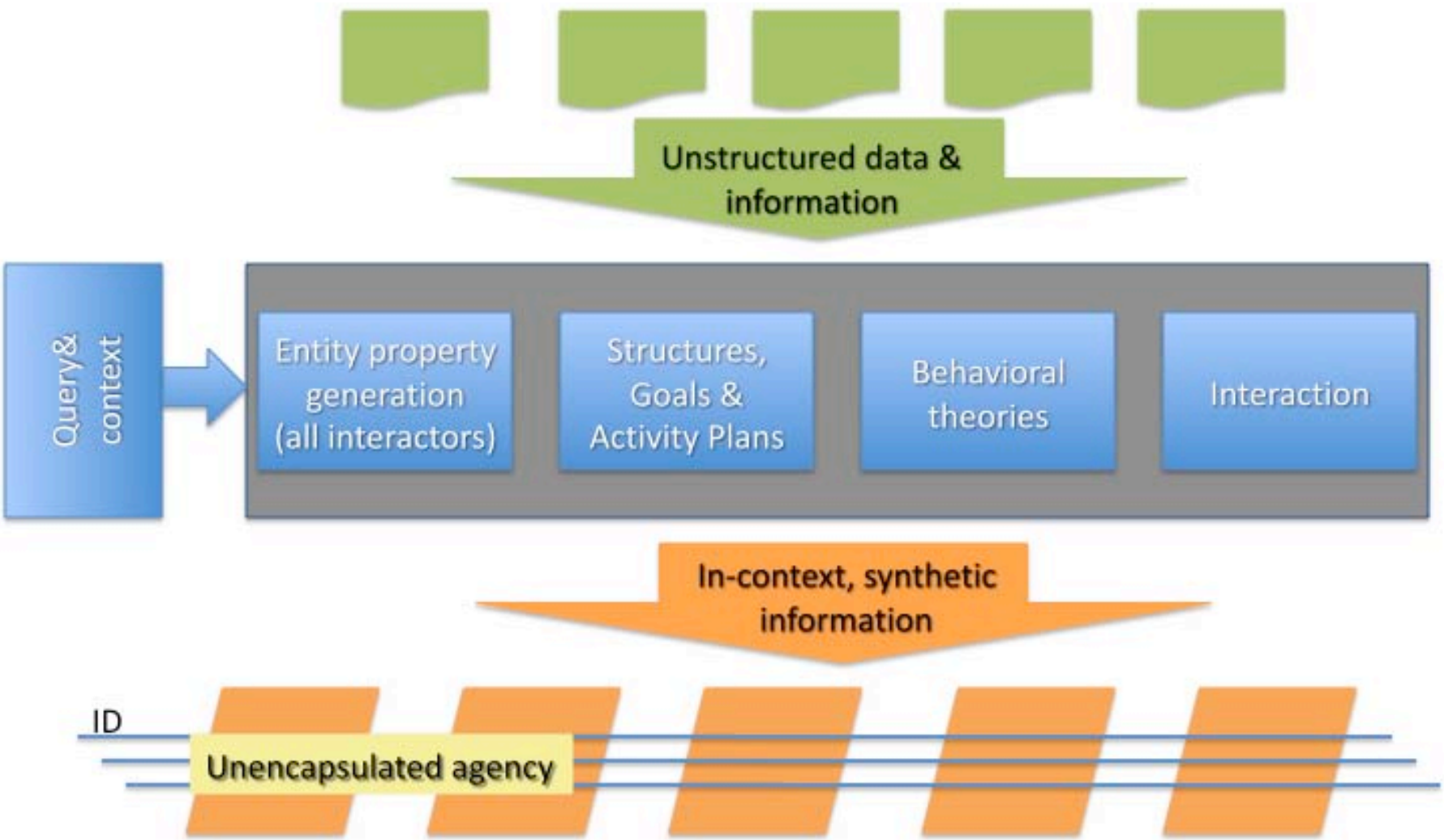




- This demonstration does not complete the question, it shows an approach
- Depends on both local interactor detail and co- evolving network topolgy

Decision & Policy Making in Complex Environments

HPC-based Synthetic Information Services Platform



Public Health Policy Planning Example

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 government cost of AV through markets?

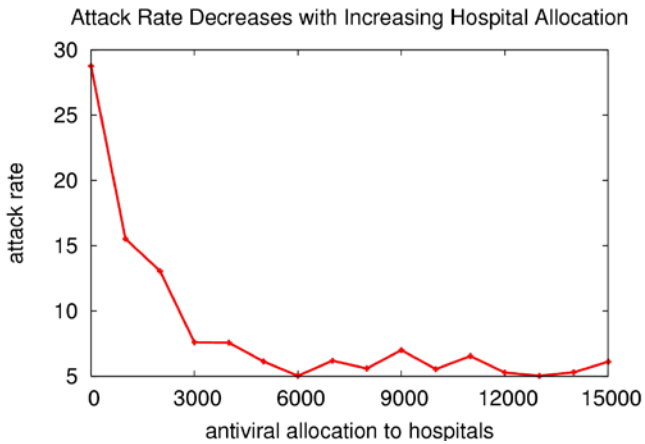
Co-Evolution Setup:

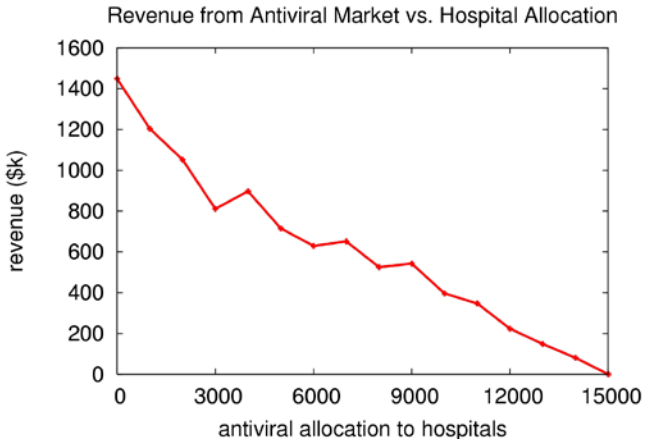
Demand → price → Inventory → likelihood of individual buying it → susceptibility → social network structure → disease dynamics → prevalence → Demand

Question: How do disease prevalence and AV demand co-evolve?

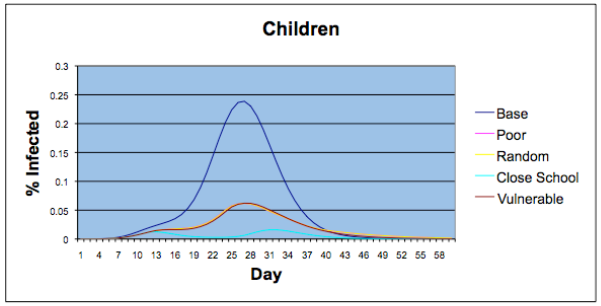
Experiment Results

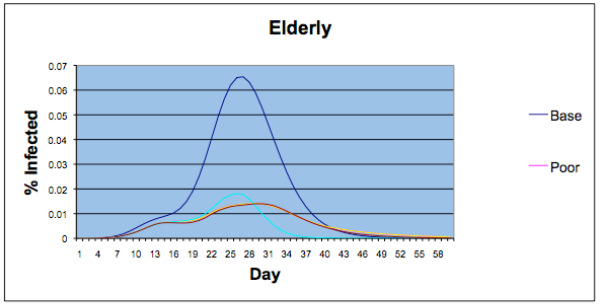
- Attack rate reaches minimum when hospital allocation of AV is 6K.
- Application of hospital allocation is upper bounded by attack rate. No need to allocate more to hospitals.
- Extra AV can go to the market. Revenue recovers cost.

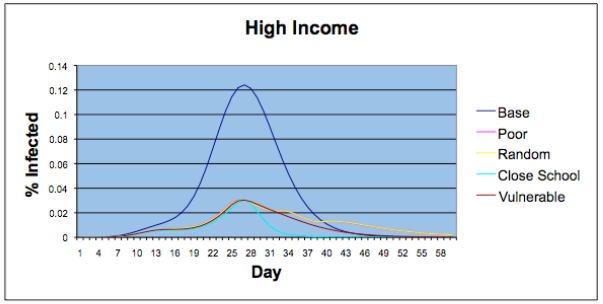


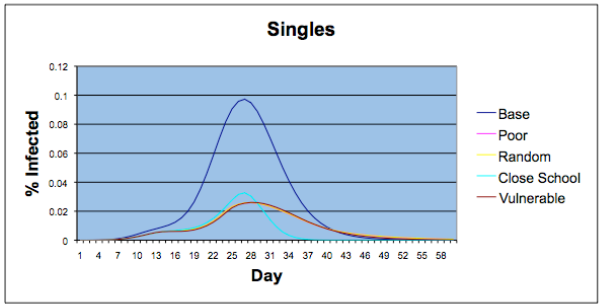


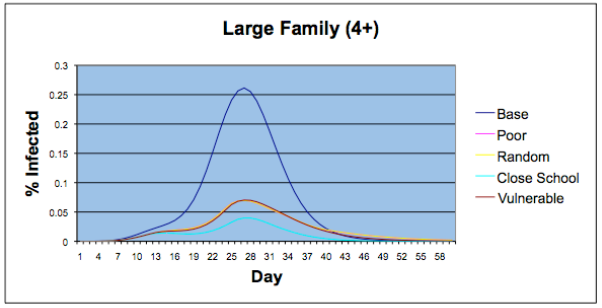
Experiment Results: Demographic Detail

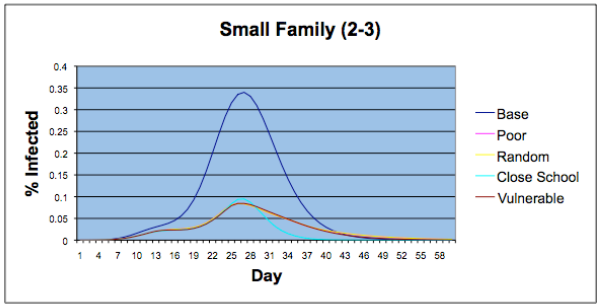


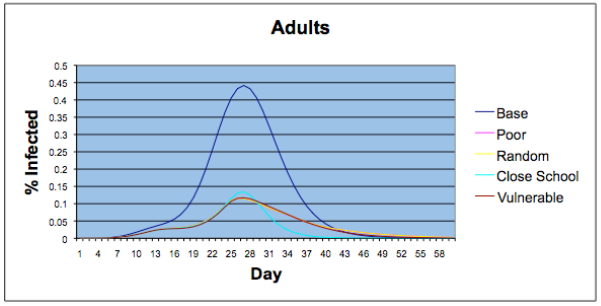


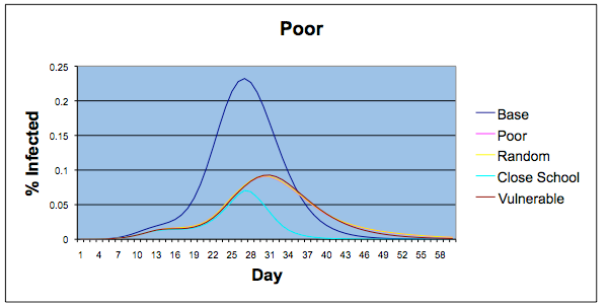


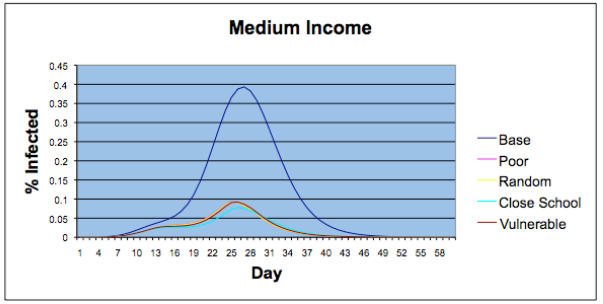












I gratefully acknowledge everyone at ndssl.vbi.vt.edu, and accept any errors in this presentation as my own. This work was sponsored by DoD, NIH, NSF and CDC.