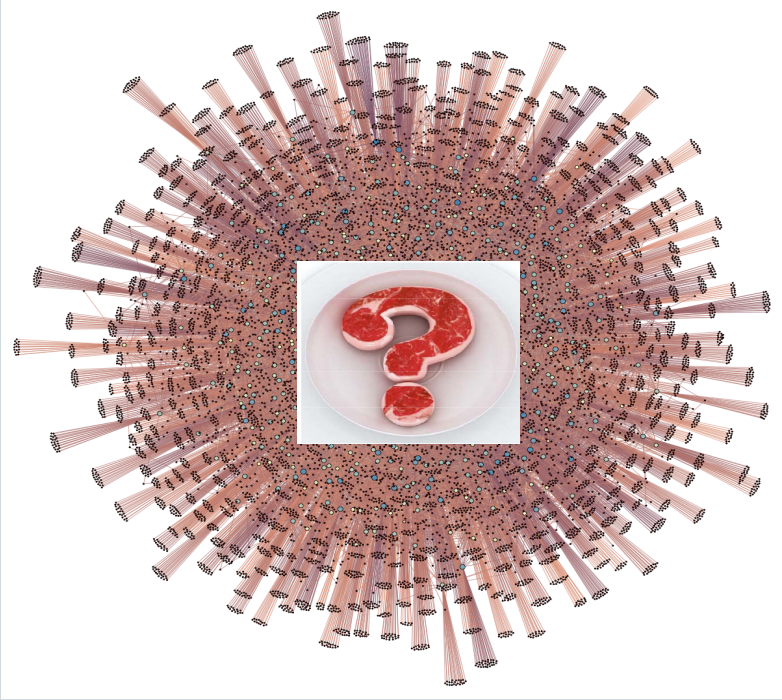
Modeling Supply Chain System Structure to Trace Sources of Food Contamination

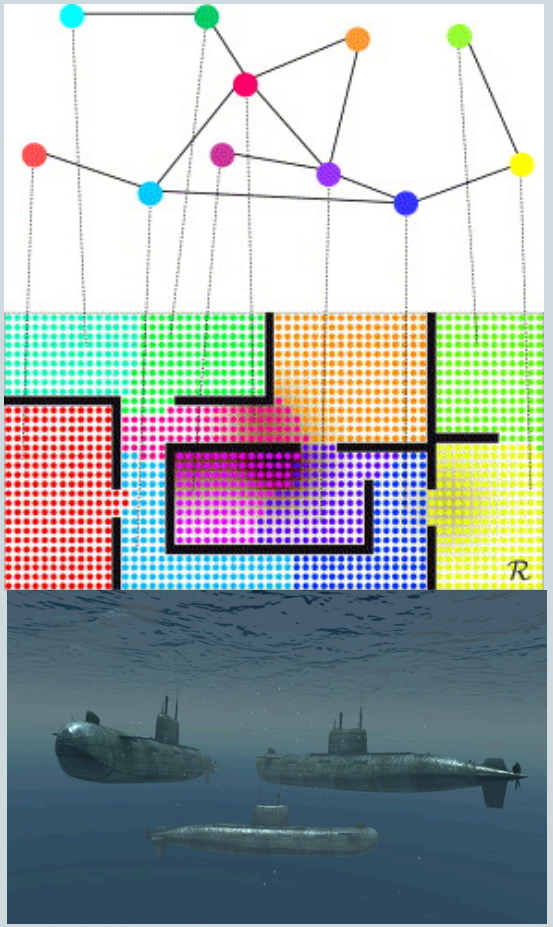


ABIGAIL HORN,
STAN FINKELSTEIN, MD,
RICHARD LARSON, PhD

Engineering Systems Division,
Massachusetts Institute of Technology (MIT)

Problem Framing: Optimal Search Theory

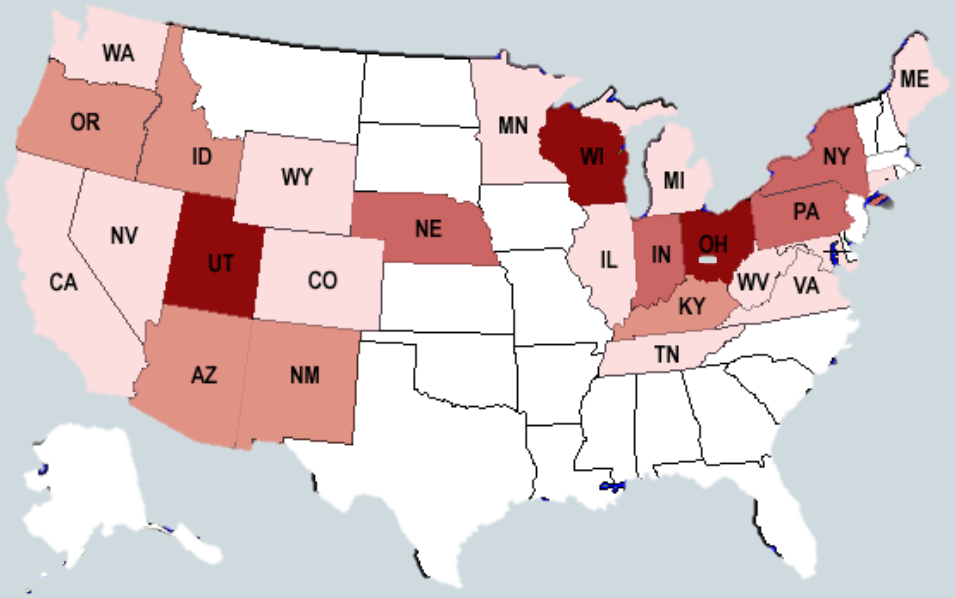
- Bernard Koopman's "Theory of Optimal Search" (Richardson, 1986).
- Anti-submarine warfare problem
  - Search over 2-D space
  - Prior probabilities
  - Bayesian updates
  - Allocated "search effort" a highly nonlinear function of the updated probabilities
- Turned around the war in the North Atlantic (Nunn, 1981).



In 2006 there was an outbreak of E. coli O157:H7 in spinach in the US

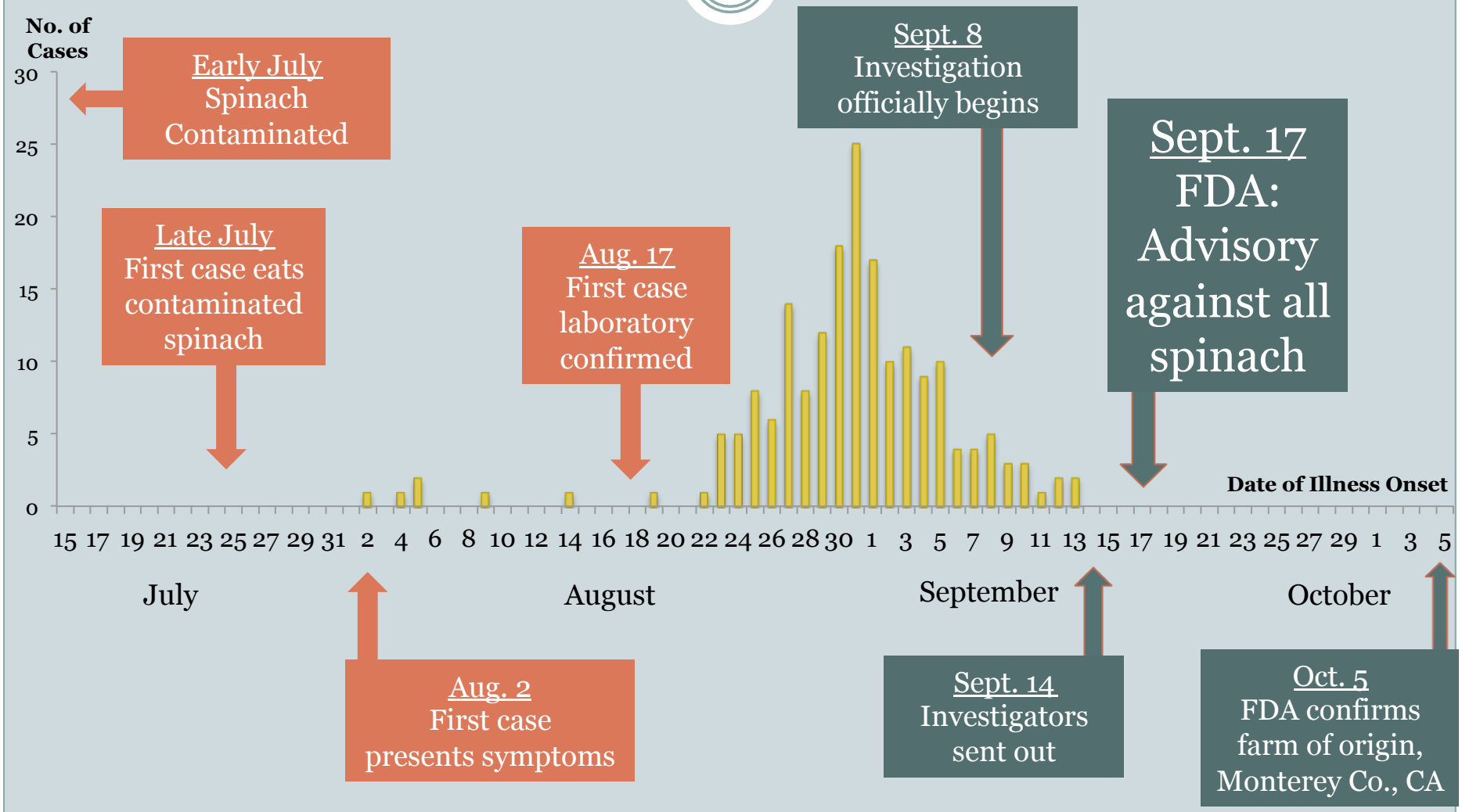
Known Impact of 2006 spinach outbreak:

- 276 Illnesses
- 102 Hospitalizations
- 3 Deaths
- 26 States with cases
- $350 million direct losses to spinach industry



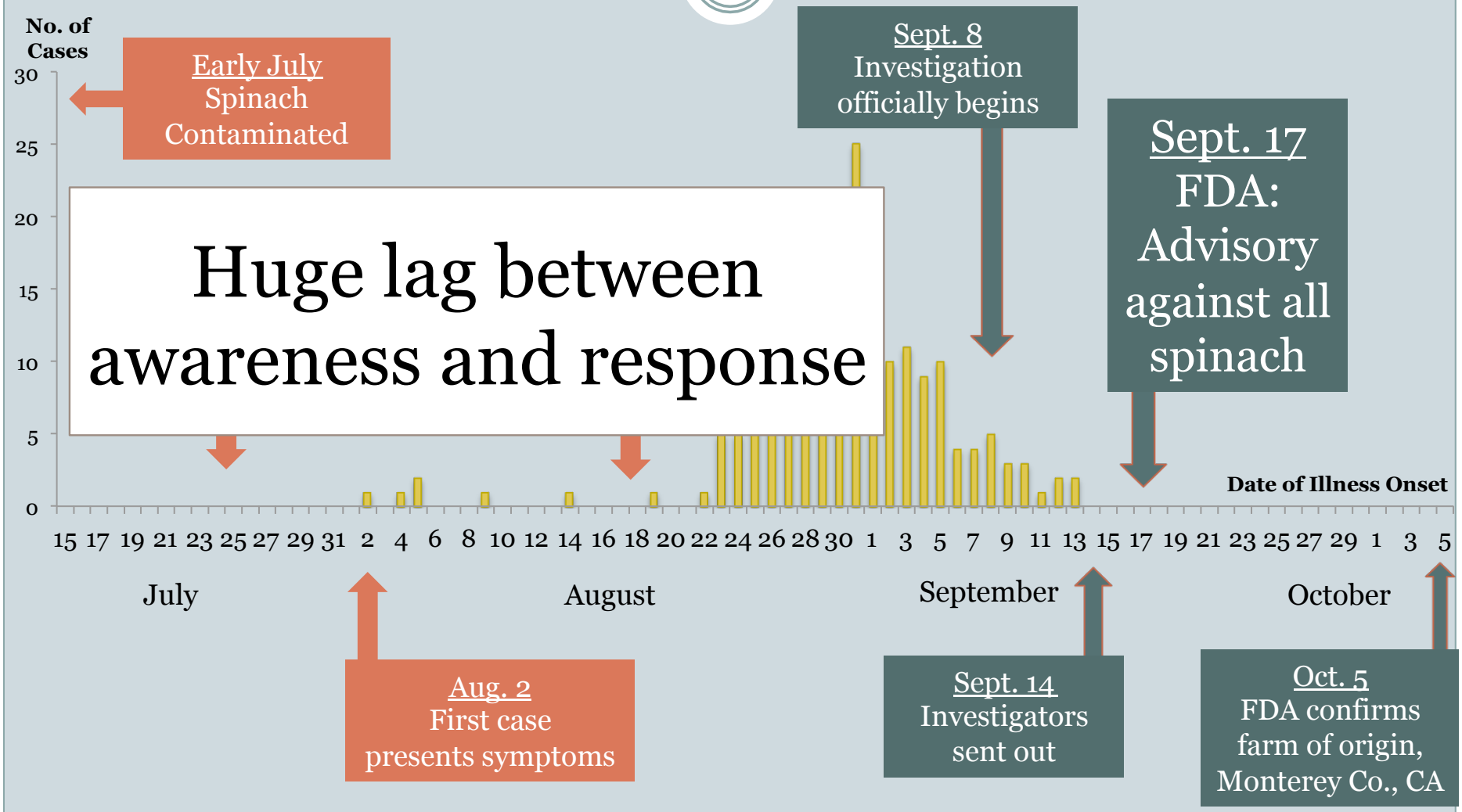
- Centers for Disease Control and Prevention (CDC) (2006a). Ongoing Multi-State Outbreak of Escherichia coli serotype O157:H7 Infections Associated with Consumption of Fresh Spinach. Morbidity and Mortality Weekly Report, 55(Dispatch); 1-2. September 26, 2006.
-California Department of Public Health (CDPH) (2007). Investigation of an Escherichia coli O157:H7 Outbreak Associated with Dole Pre-Packaged Spinach, Final Report prepared by the California Food Emergency Response Team. March 21, 2007. http://www.cdph.ca.gov

Could damage have been prevented?



California Department of Public Health (CDPH) (2007). Investigation of an Escherichia coli O157:H7 Outbreak Associated with Dole Pre-Packaged Spinach, Final Report prepared by the California Food Emergency Response Team. March 21, 2007. http://www.cdph.ca.gov.

Could damage have been prevented?



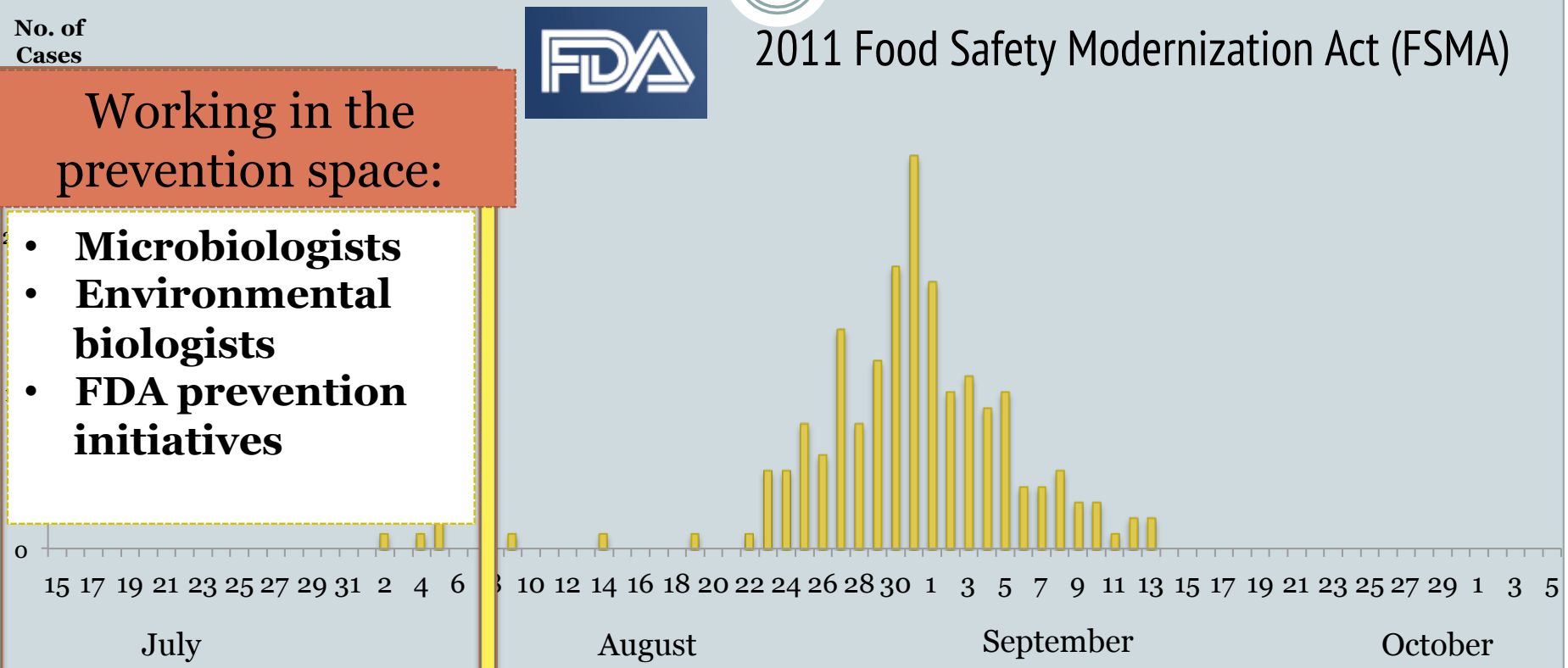
California Department of Public Health (CDPH) (2007). Investigation of an Escherichia coli O157:H7 Outbreak Associated with Dole Pre-Packaged Spinach, Final Report prepared by the California Food Emergency Response Team. March 21, 2007. http://www.cdph.ca.gov.

Outbreak Prevention

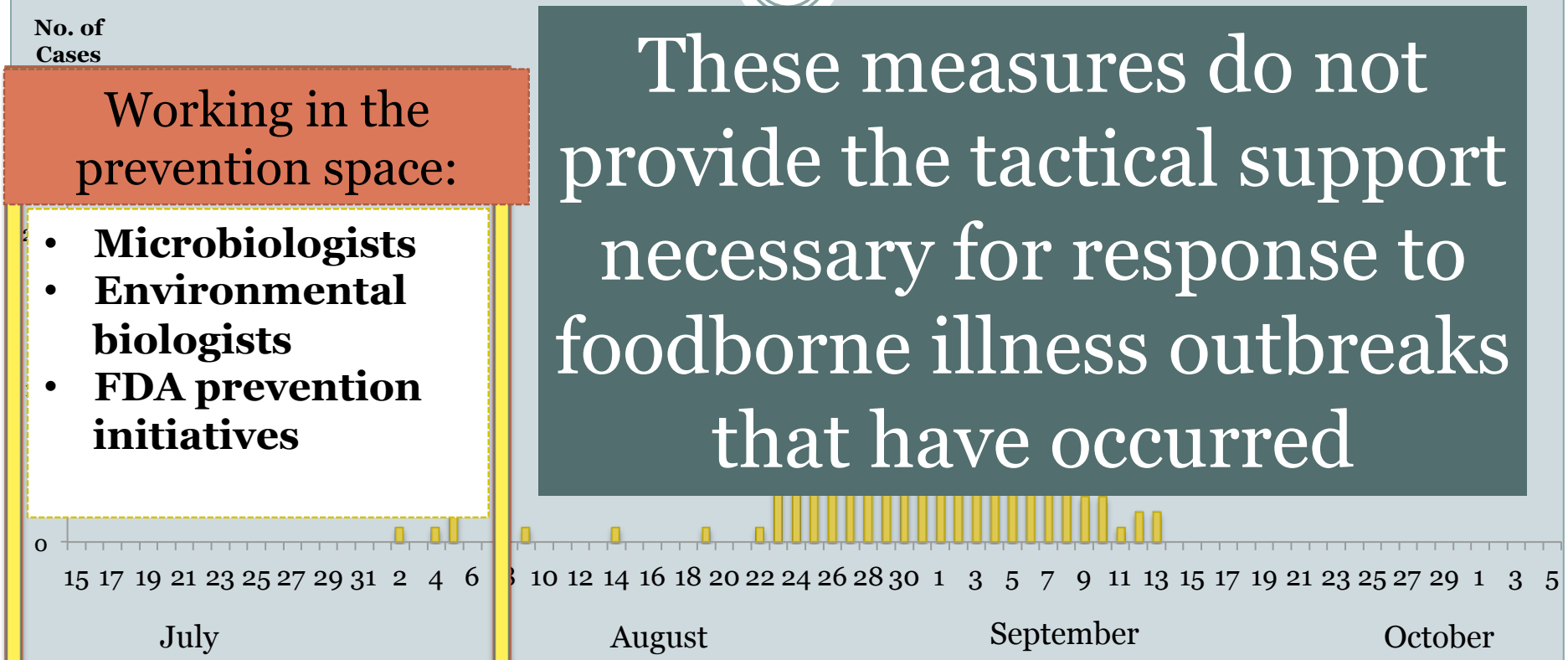
2011 Food Safety Modernization Act (FSMA)

Working in the prevention space:

- **Microbiologists**
- **Environmental biologists**
- **FDA prevention initiatives**



Outbreak Prevention



Working in the prevention space:

- **Microbiologists**
- **Environmental biologists**
- **FDA prevention initiatives**

These measures do not provide the tactical support necessary for response to foodborne illness outbreaks that have occurred

Annual Impact of Foodborne Disease Outbreaks

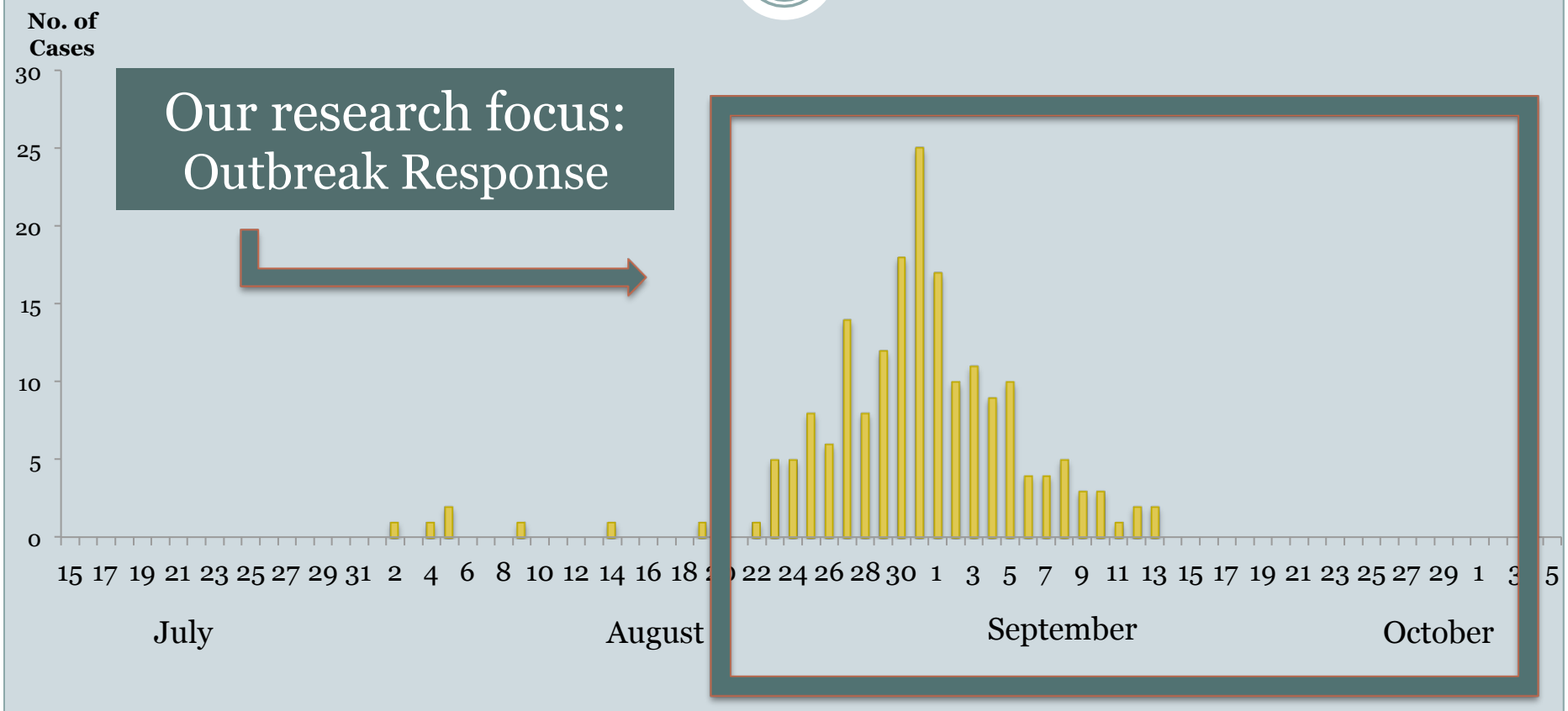
Despite efforts at prevention
the impact of foodborne disease outbreaks remains high:

Every year in the US…

- 48 million illnesses, 128,000 hospitalizations, 3000 deaths
- $77 billion in healthcare costs
- **>65% of identified foodborne illness outbreaks UNSOLVED**

- Osterholm, MT. Foodborne Disease in 2011 — The Rest of the Story. N Engl J Med 2011; 364:889-891, March 10, 2011.
- Scharff, R. (2009). Health-related costs from food borne illness in the United States. Retrieved from http://www.producesafetyproject.org
- Jennifer B. Nuzzo, Samuel B. Wollner, Ryan C. Morhard, Tara Kirk Sell, Anita J. Cicero, Thomas V. Inglesby. (2013). When Good Food Goes Bad: Strengthening the US Response to Foodborne Disease Outbreak. Final Report: Center for Biosecurity of UPMC.

Tactical Response to Outbreaks



Outbreak Response: Opportunities for Improvement

There are several opportunities for improving the process of outbreak response that can have positive and meaningful impacts on public health:

- Recognizing when a foodborne illness outbreak is occurring – **outbreak identification**
- Identifying the food and location source of contamination – **outbreak traceback**

Outbreak Identification

Government surveillance systems

- **PulseNet** – performs "DNA fingerprinting" of pathogens
- FoodCORE
- FoodNET / OutbreakNet
- VetNET

Digital disease detection

- Mining information from social networks e.g. Yelp, Twitter
- Active research area, early models show success in detecting cases and identifying outbreaks

Centers for Disease Control and Prevention (CDC) (2006b). Timeline for Reporting of E. coli Cases. September 19, 2006. http://www.cdc.gov/ecoli/reportingtimeline.htm

Toner ES, Nuzzo JB, Watson M, et al. Biosurveillance where it happens: state and local capabilities and needs. *Biosecur Bioterror*. 2011 Dec;9(4):321-330.

Traceback Investigation

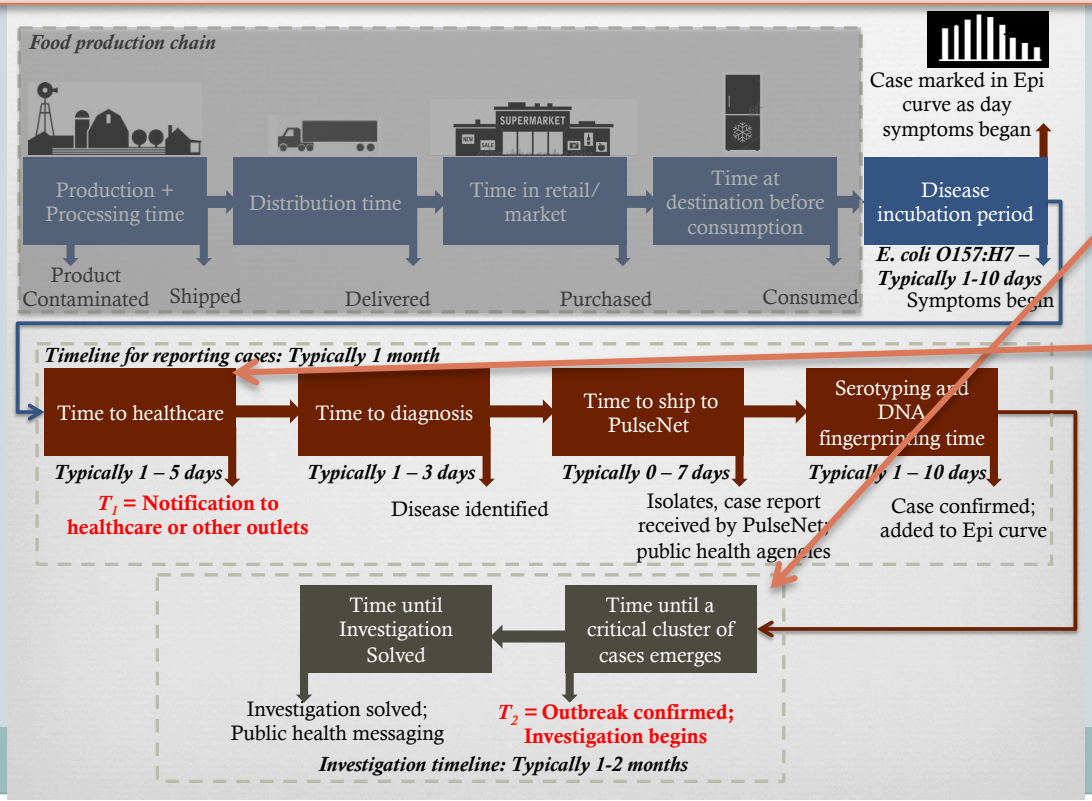
Current investigative methods are slow, resource intensive, and often unsuccessful

- Rely on in-person data collection
- Do not leverage modern data and analytics
- Limited by delays between initial case reports and launching of the investigation…

TB Investigation: Limited by Delays

The traceback investigation does not begin until a "critical cluster" of cases have been confirmed as linked to a single outbreak strain by PulseNet.

The pathway from contamination to illness, set of illnesses to confirmed outbreak, and confirmed outbreak to resolved investigation



Investigation begins only here

But it could begin here

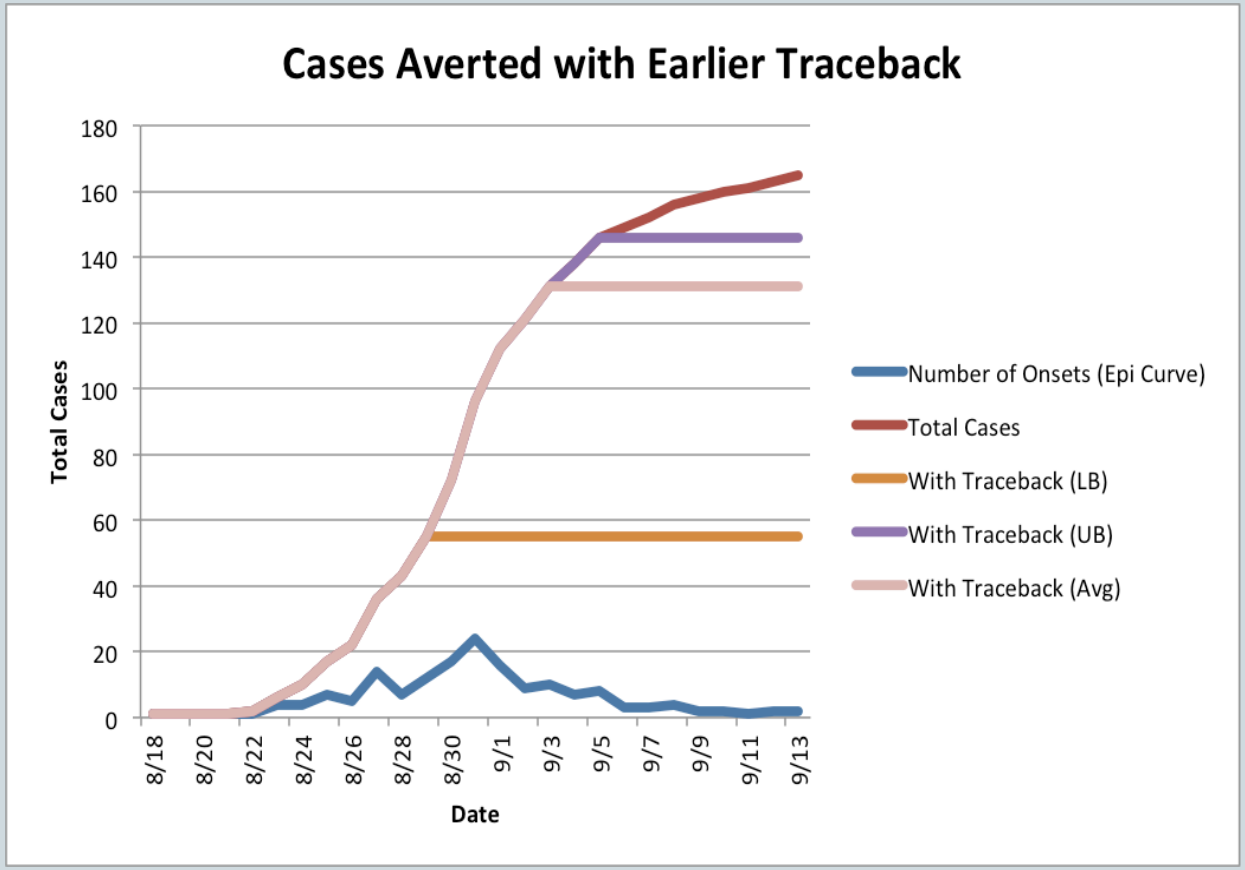
TB Investigation: Opportunities for Improvement

We advocate leveraging initial case diagnoses to enable an earlier, "exploratory" investigation of convergent sources of contamination.

↳ In advance of sending out investigators to collect records documenting product pathways
↳ Through a low-cost, low-commitment computer model, which identifies and rules out possible sources

This "exploratory" investigation would help to prioritize leads early on, enabling a faster resolution of the investigation.

TB Investigation: Time Saved = Illnesses Averted



Confirmed Illnesses Potentially Avoided with Earlier Traceback

Lower Bound: **110**
Upper Bound: **19**
Average: **34**

Illnesses potentially avoided through earlier traceback in the 2006 E. coli in spinach outbreak

Bayesian Updating Network Approach

How can the process of tracing the source of an outbreak be sped up?

Prior information

From past outbreaks, causal factors, current consumption, etc.

+

Investigations occur over the supply chain →

Leverage what is known about **network structure**

+

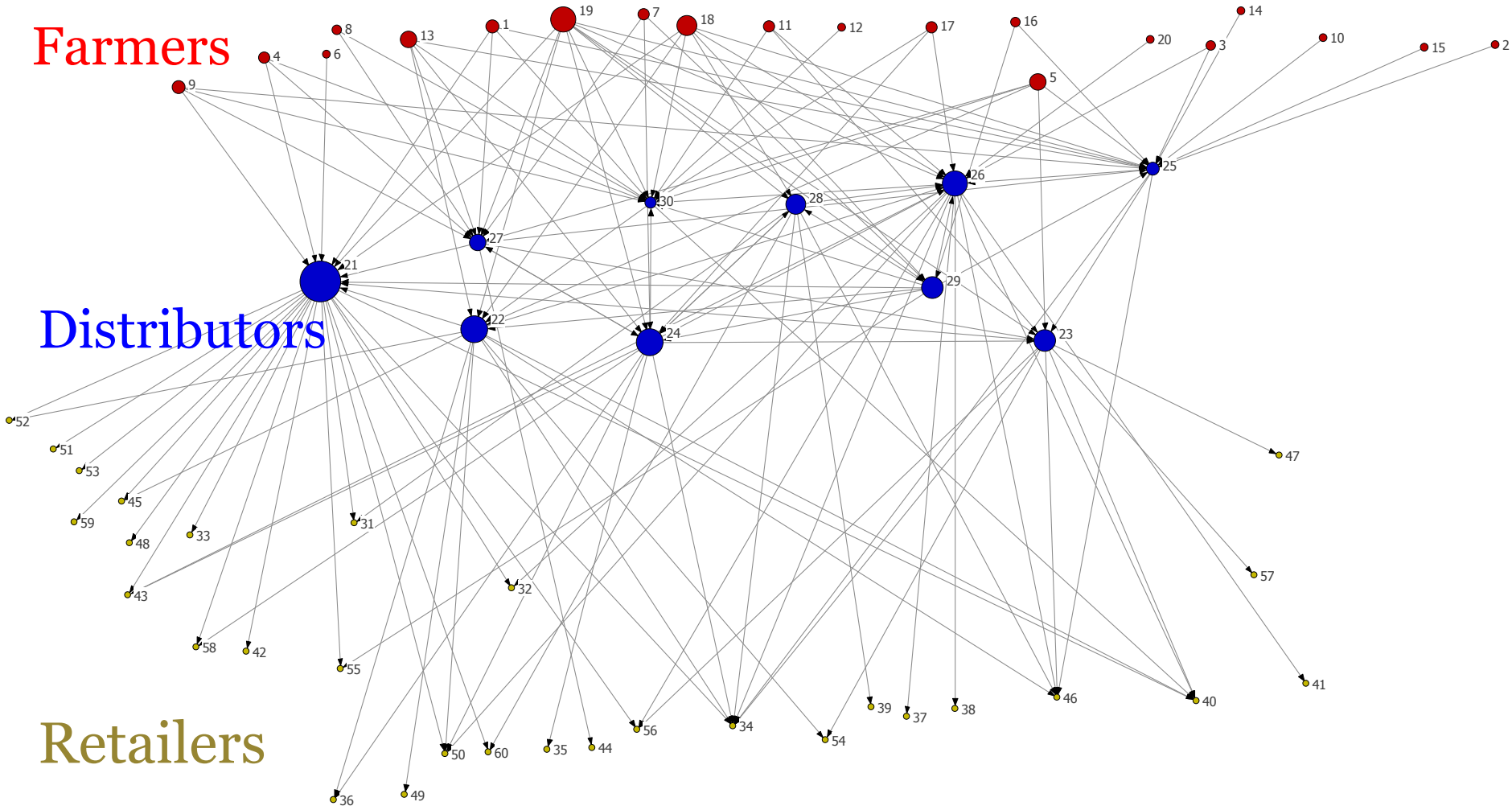
Network is dynamic, imperfectly understood; search ongoing →

Method should allow for **dynamic updating**

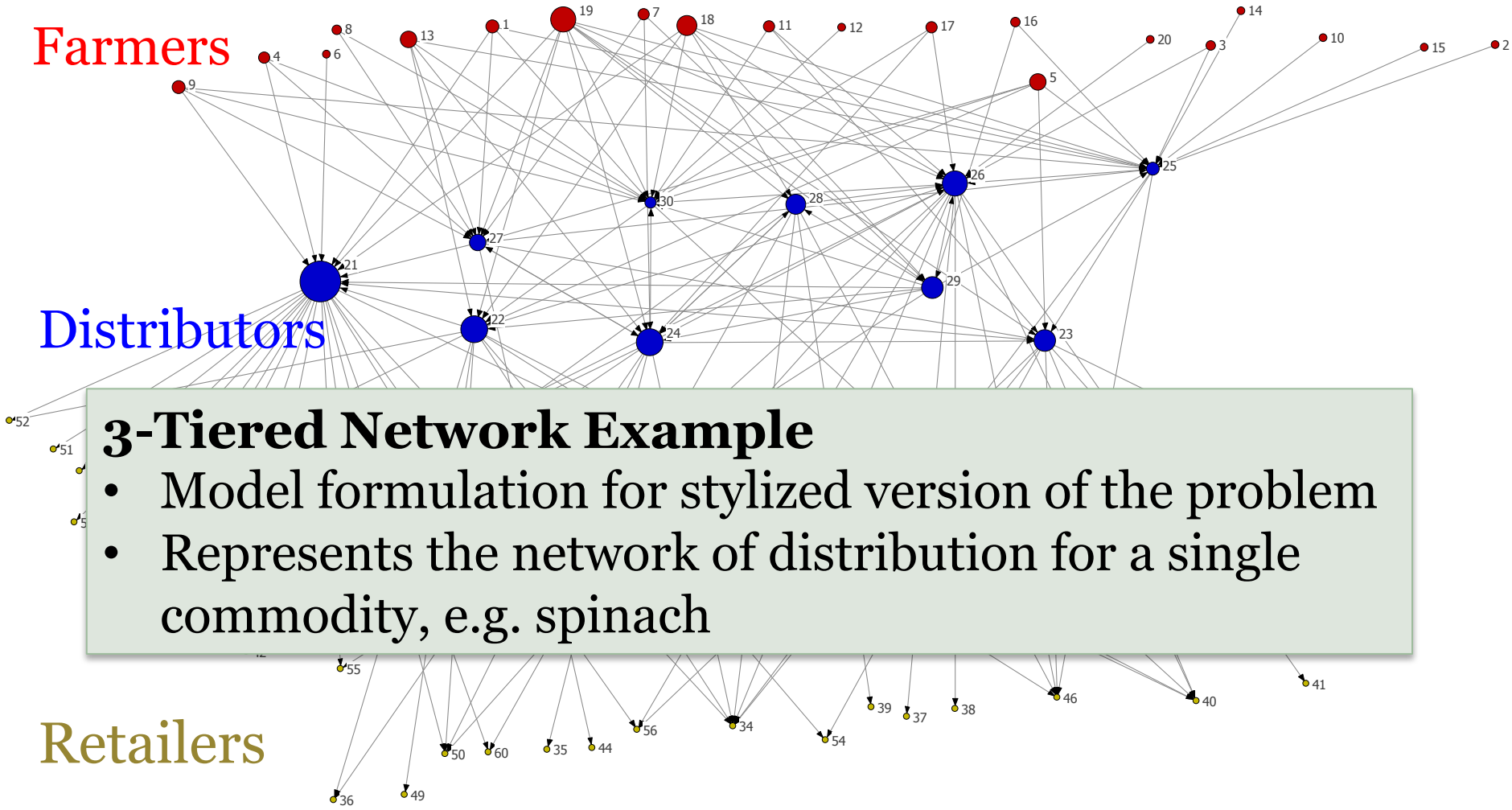
=

BAYESIAN UPDATING NETWORK APPROACH

Modeling Framework – Stylized Example



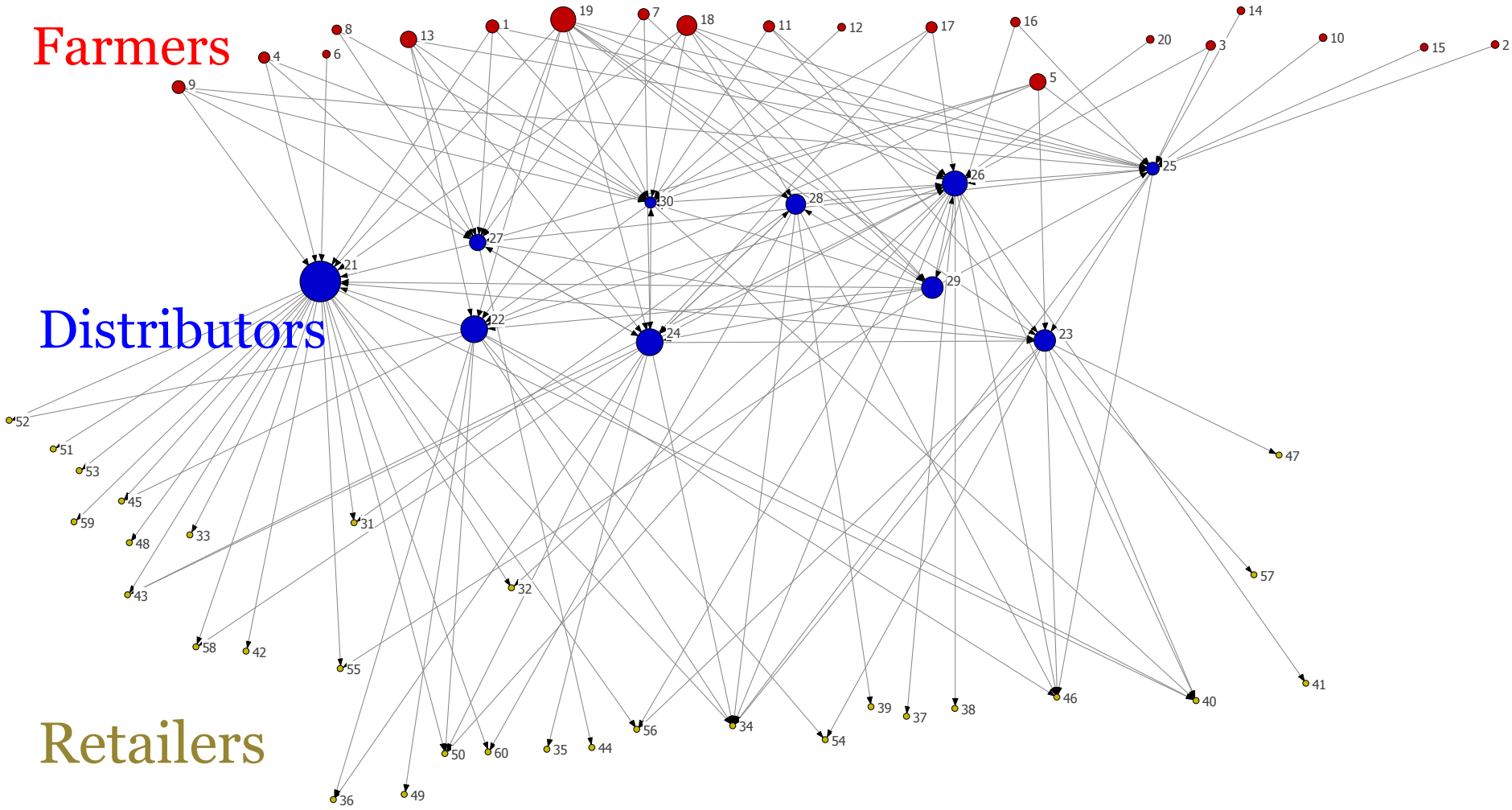
Modeling Framework – Stylized Example



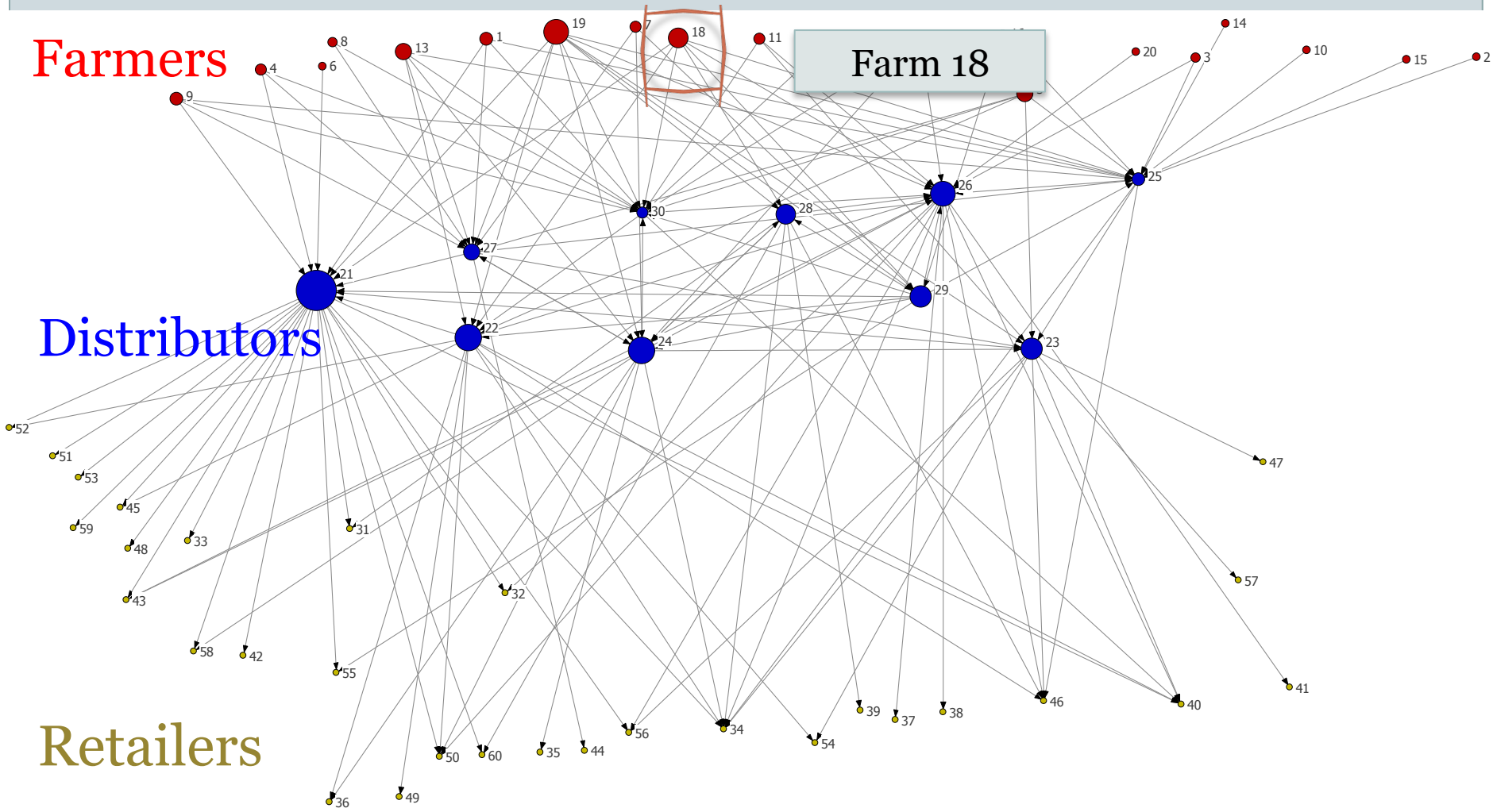
3-Tiered Network Example

- Model formulation for stylized version of the problem
- Represents the network of distribution for a single commodity, e.g. spinach

Modeling Framework – Stylized Example

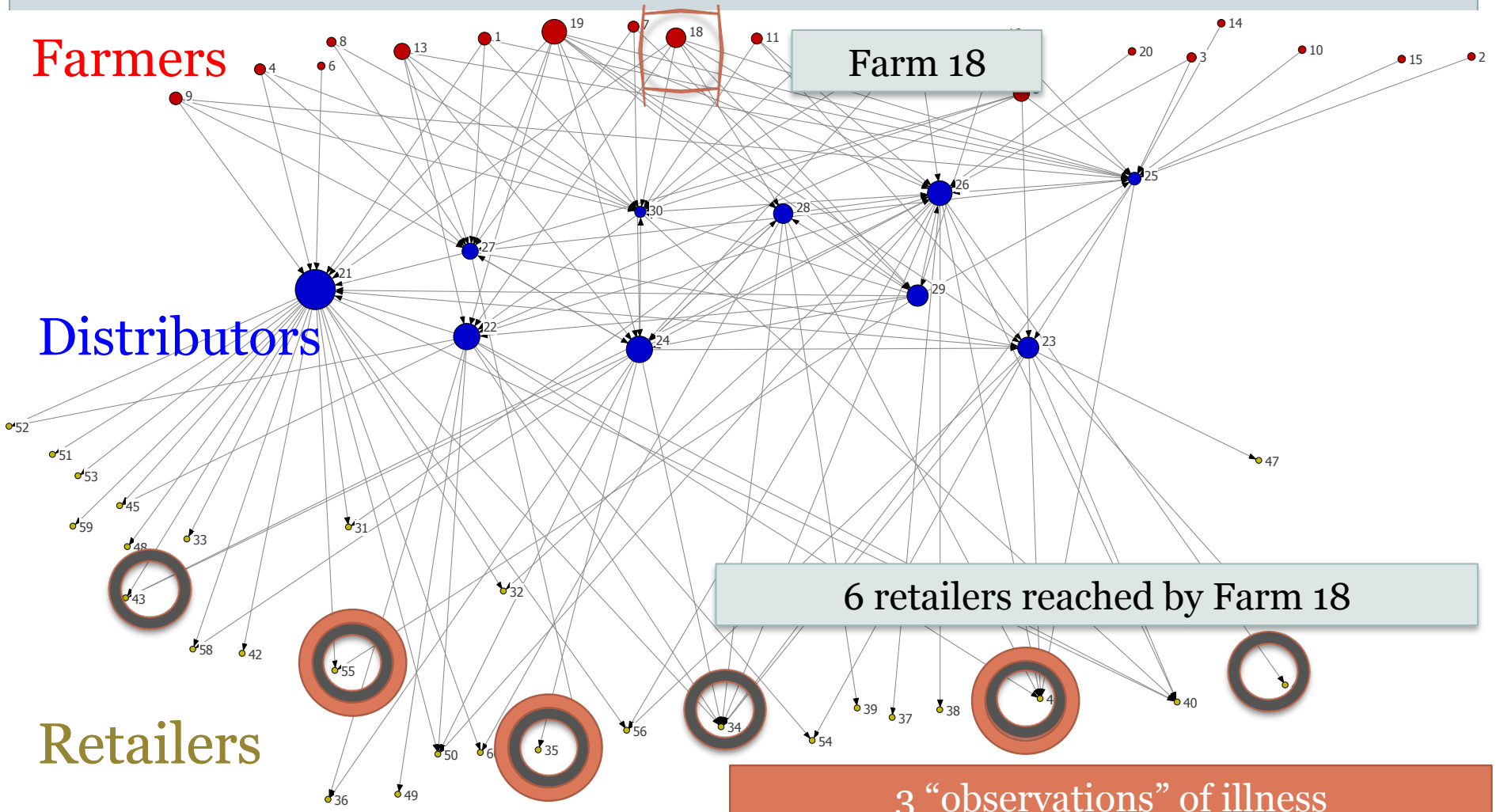


Modeling Framework – Stylized Example

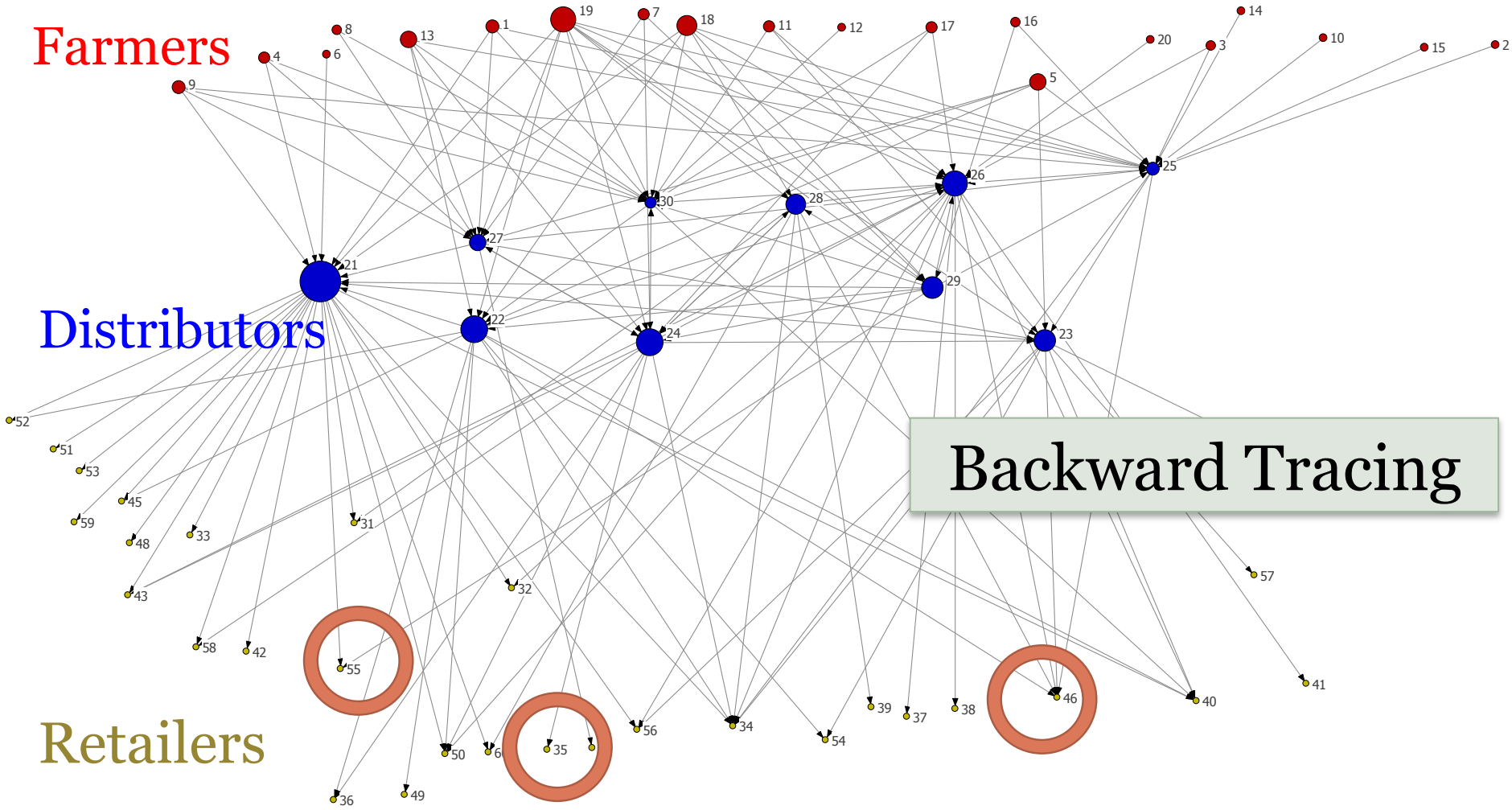


Modeling Framework – Stylized Example

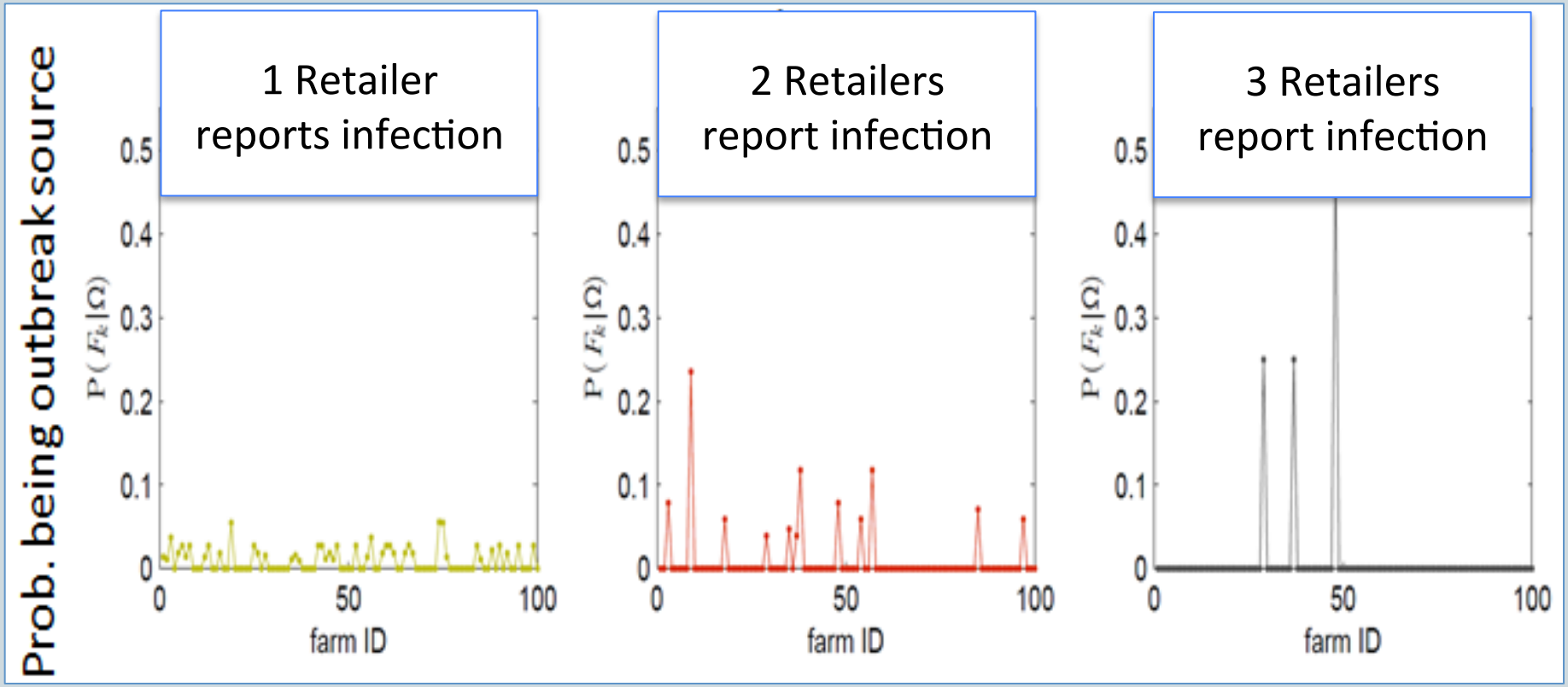
Modeling Framework – Stylized Example



Modeling Framework – Stylized Example



Accuracy improves as case count increases



Research Approach: Increasingly Complex Decision Space

1. Prior Probabilities about outbreak source location

Identify locations that could be the source of an ongoing outbreak and determine the prior probability that any of these locations is the source

2. Analytical models using simple, stylized network structures

Come up with exact results and algorithms that lead to new, general insights into the relationship between network structure and traceback accuracy

3. Simulation models that include true system complexity

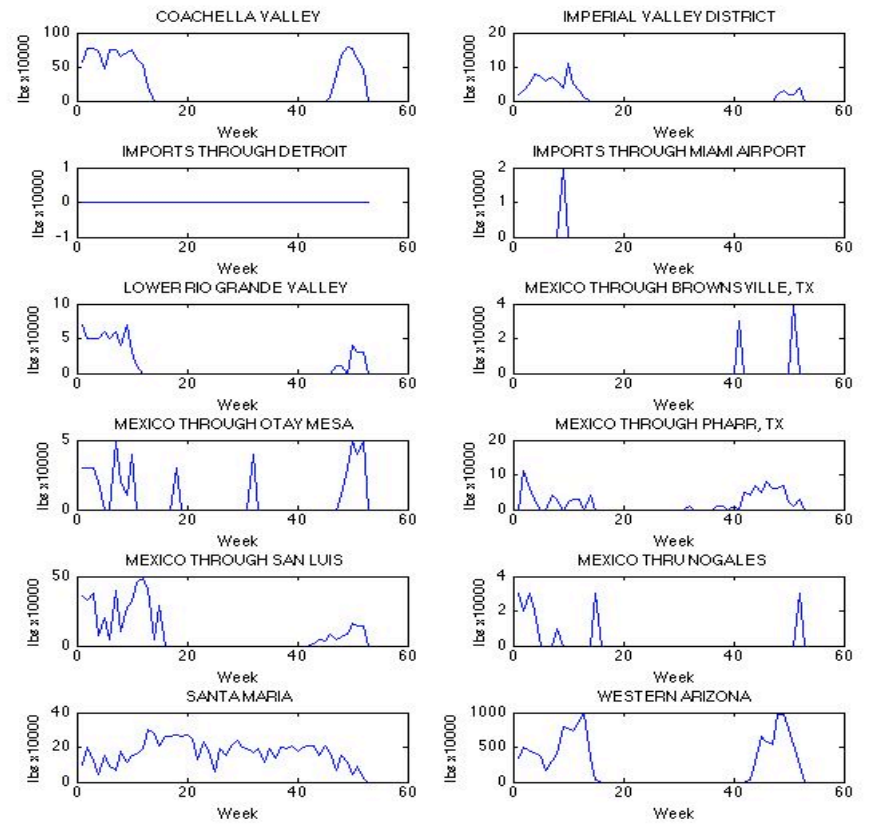
Develop a Monte Carlo simulation framework that incorporates many more of the real-world complexities of the outbreak generation, spreading, reporting, and traceback process

1. Bayesian Prior Probabilities

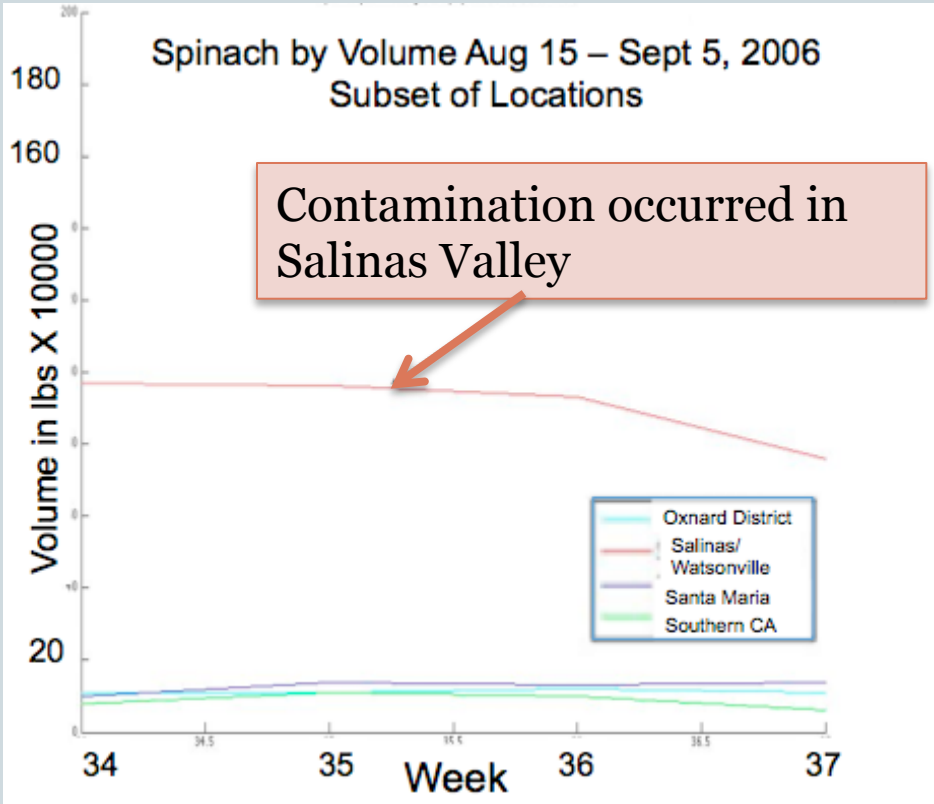
A Bayesian prior probability model has been developed to identify a set of feasible outbreak locations, and to assign relative risks according to the following three risk dimensions:

1) **Current consumption:** likelihood of exposure given relative consumption rates
2) **Contamination-causing events:** likelihood of observed contamination-causing events to have caused all outbreak illnesses
3) **Epidemiological association:** strength of the historic epidemiological association between outbreaks in a location and a specific pathogen-commodity pair

1. Bayesian Prior Probabilities

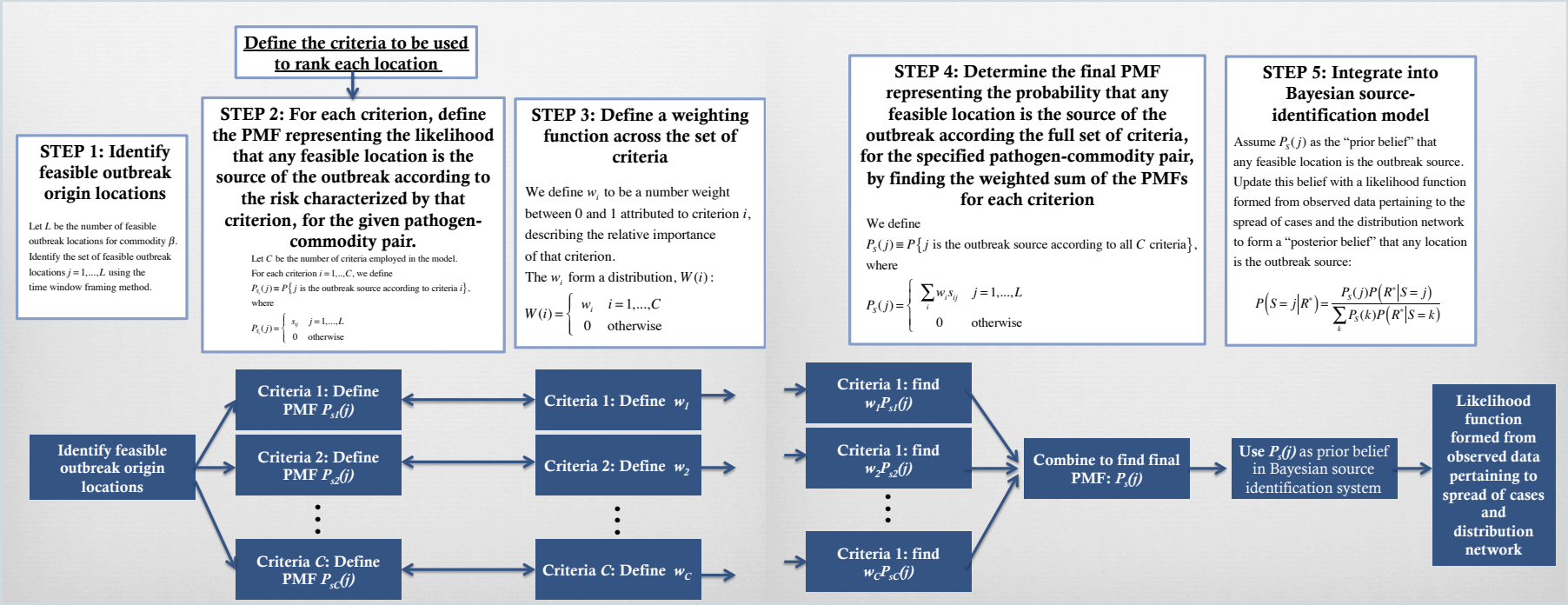


Shipping records for volume of spinach shipped from all (12) districts of origin during 2006 (Source: USDA AMS)



Volume of spinach for ***possible*** origin districts during relevant dates (August 15 – September 6th, 2006).

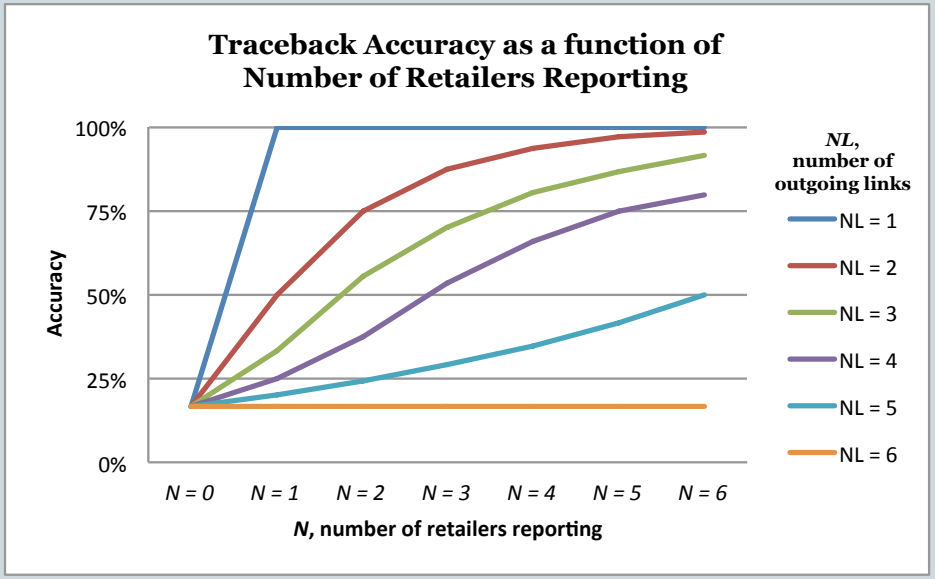
Method for Assigning Prior Probabilities

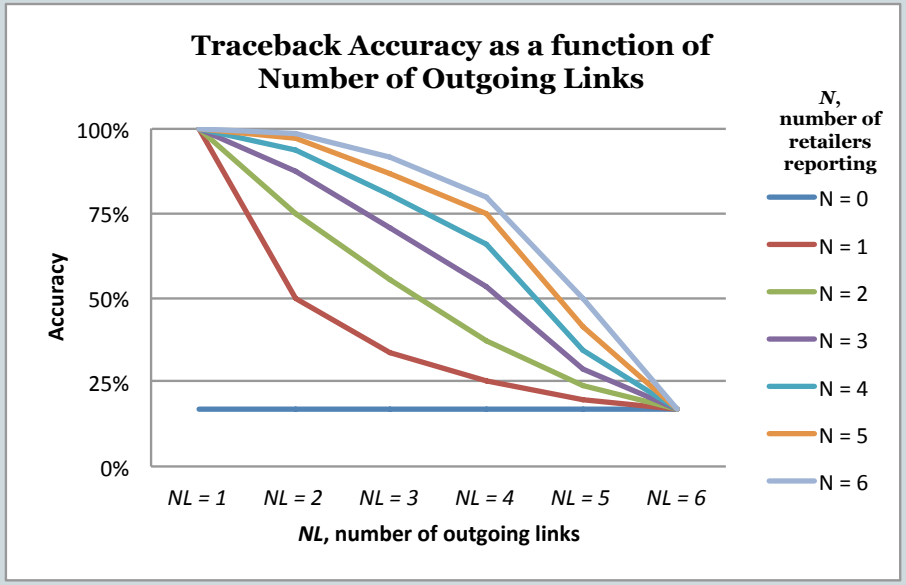


2. Analytical Models

Using simple, stylized network structures:

→ Develop expressions to explore the relationship between network parameters and traceback accuracy, e.g.:





→ The math provides key insights to relationships such as **tradeoffs**

3. Simulation on Realistic Modeling Structures

Use a Monte Carlo simulation framework to represent the stochastic, dynamic, and imperfectly understood nature of real food distribution networks:

- Vary the network variables to explore multiple **connectivity patterns**
- Vary the outbreak variables to explore the effect of the **size**, **origin**, **case count**, and **timing** of an outbreak

Test methodology using simulated network structures

- Determine traceback accuracy under various conditions
- What parameters have the biggest influence on the traceback process and how can we influence them through policy changes?
  - ↳ Compare **hypothetical policies** for improving traceability, **determine which could have the greatest impact**

Improve Accuracy and Speed of Tracebacks

Tactically: Develop *implementable* set of procedures to guide investigation and control measures during outbreaks

- Allocation of search effort
- Where additional data collection is warranted
- When to issue public service announcements

Strategically: Prioritize hypothetical policies for improving outbreak response process, e.g.

- Strengthen risk-based prior probability through standardized data collection
- Proactive mapping of network structures for high-risk foods
- Improved surveillance / earlier detection of cases
- Holding samples at each farm / distribution center

Jennifer B. Nuzzo et al (2013). When Good Food Goes Bad: Strengthening the US Response to Foodborne Disease Outbreak. Final Report: UPMC.

Expected Contributions

Academic and practical contribution is to improve accuracy and speed of tracebacks

"Any measure that will help to determine where we should focus our attention and give leads on the investigation would have a lot of application and utility for public health. Messaging could be more targeted because we would be able to narrow down more quickly where the product is not coming from…This could really make a difference early on!"

–– S. McGarry, Foodborne Outbreak Coordinator at FDA Headquarters, Personal communication, December 20, 2012

Next Steps

The current project phase ends in March. Deliverables we expect by then:

- A prototype simulation model representing the outbreak generation, spreading, reporting, and traceback system;
- A detailed list of recommended policies/procedures for improving outbreak response;
- By June: Journal articles on our framework, theory, modeling work and results, and culminating paper explaining the integrated system and potential impact on public health (to submit to *Service Science*, *Value in Health*, and *Frontiers in PHSSR*);
- Policy brief describing how this methodology might be implemented by the FDA and local public health agencies, and discussing its potential to improve the traceback process

Post-project dissemination activities include:

- Packaging the methodology into a shovel-ready predictive modeling tool for use by the FDA and public health agencies; working with officials to reality-test the tool
- Holding a workshop for public health officials to demonstrate the tool
- Developing an interactive visualization of the tool and results to be used for educational or demonstrational purposes

Acknowledgements

Robert Wood Johnson Foundation

PUBLIC HEALTH | SERVICES & SYSTEMS RESEARCH
PRACTICE-BASED RESEARCH NETWORKS

Thank you!

Questions?

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division