Finding the Most Vital Arc in a Network: Lessons Learned from Infrastructure Defense

David Alderson, PhD
Operations Research Department
Naval Postgraduate School
dlalders@nps.edu

Graph Exploitation Workshop

MIT Lincoln Labs

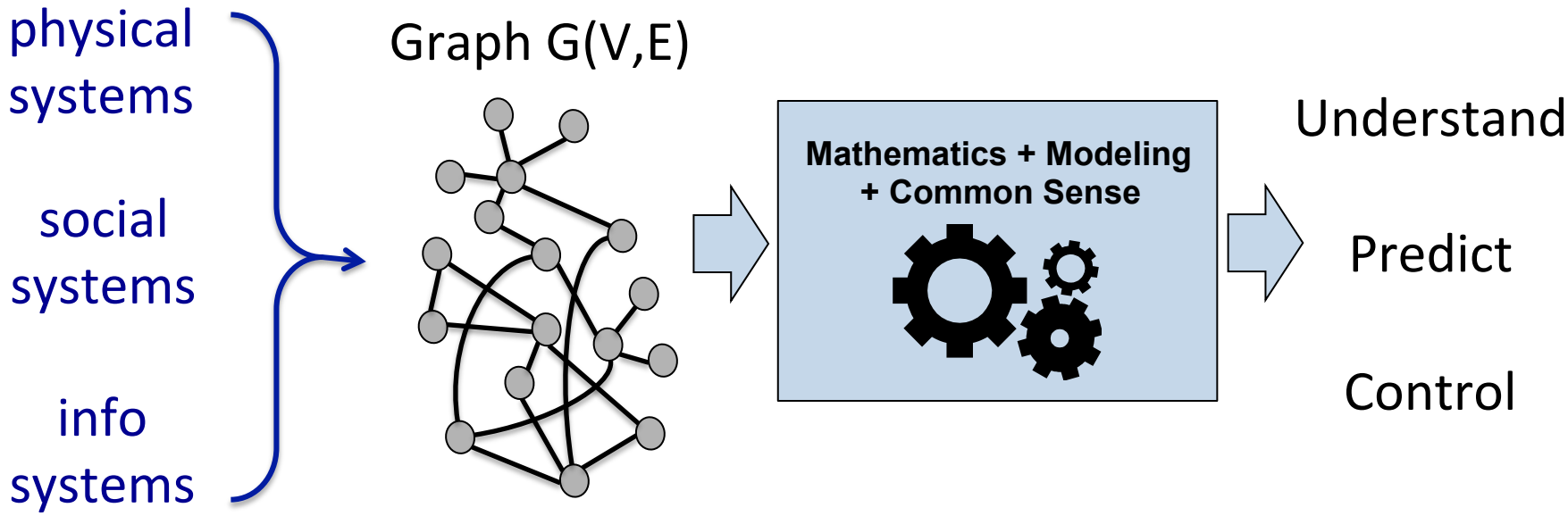
August 2011

Acknowledgments

- NPS Faculty: Jerry Brown, Matt Carlyle, Javier Salmerón, Kevin Wood
- NPS Students: Barkley, de la Cruz, Derosier, Dixon, Langford, McPherson, Nicholas, Sanchez, Yuhas
- Caltech: John Doyle, Lun Li, Steven Low
- AT&T-Research: Walter Willinger

- Supported by ONR, AFOSR, DTRA

Graph Exploitation

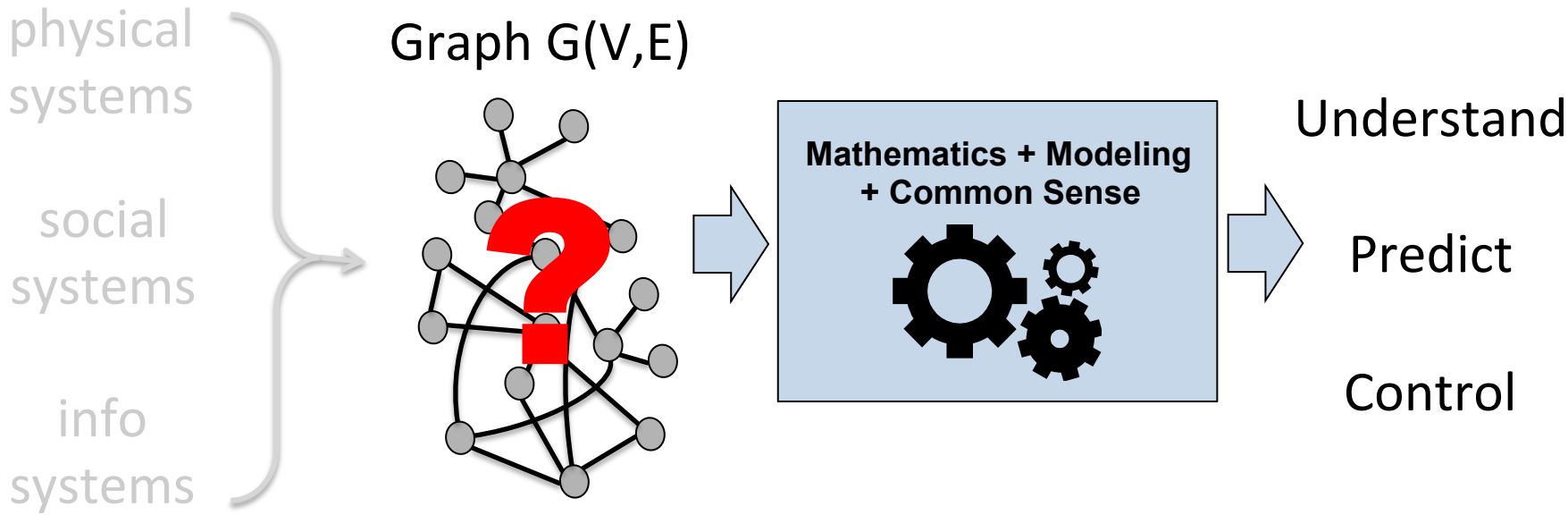


- We have a {system, artifact} with an underlying graph structure
- Much of the focus on the following:

  "Given a graph G(V,E), what can we say about it?"

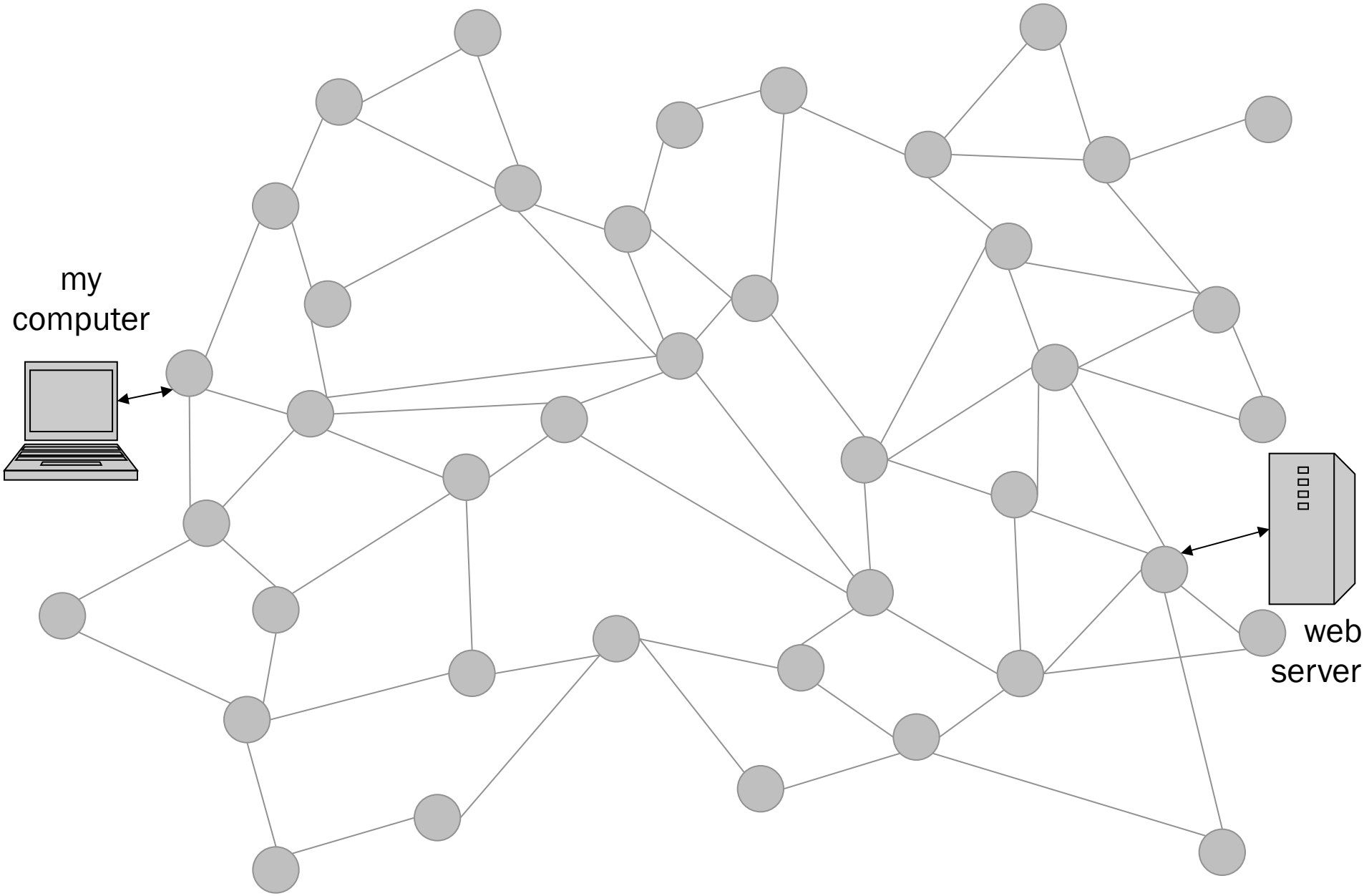
  Most commonly, this assumes that we know G(V,E)!

Pitfall #1: Does the graph represent ground truth?

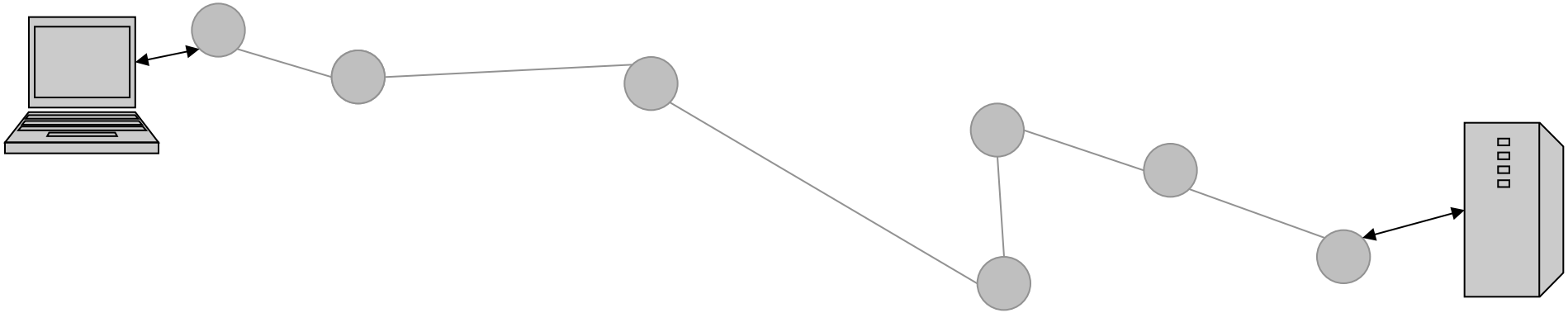


- How was the graph data collected?
- Are there biases, systematic errors, or idiosyncrasies, in the measurement data?
- What types of analyses can the graph data support?
- How does uncertainty in the underlying graph data affect the analyses that you perform?

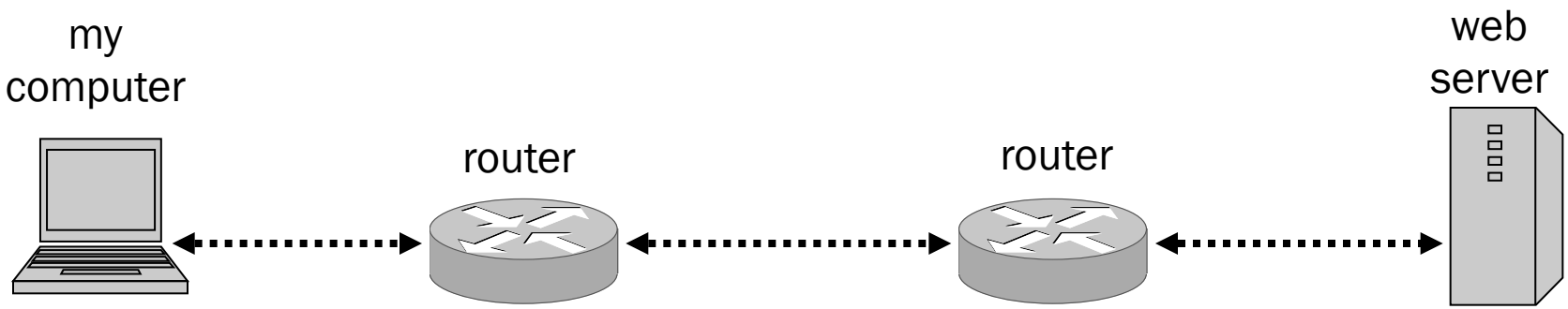
The Internet



The Internet



The Internet

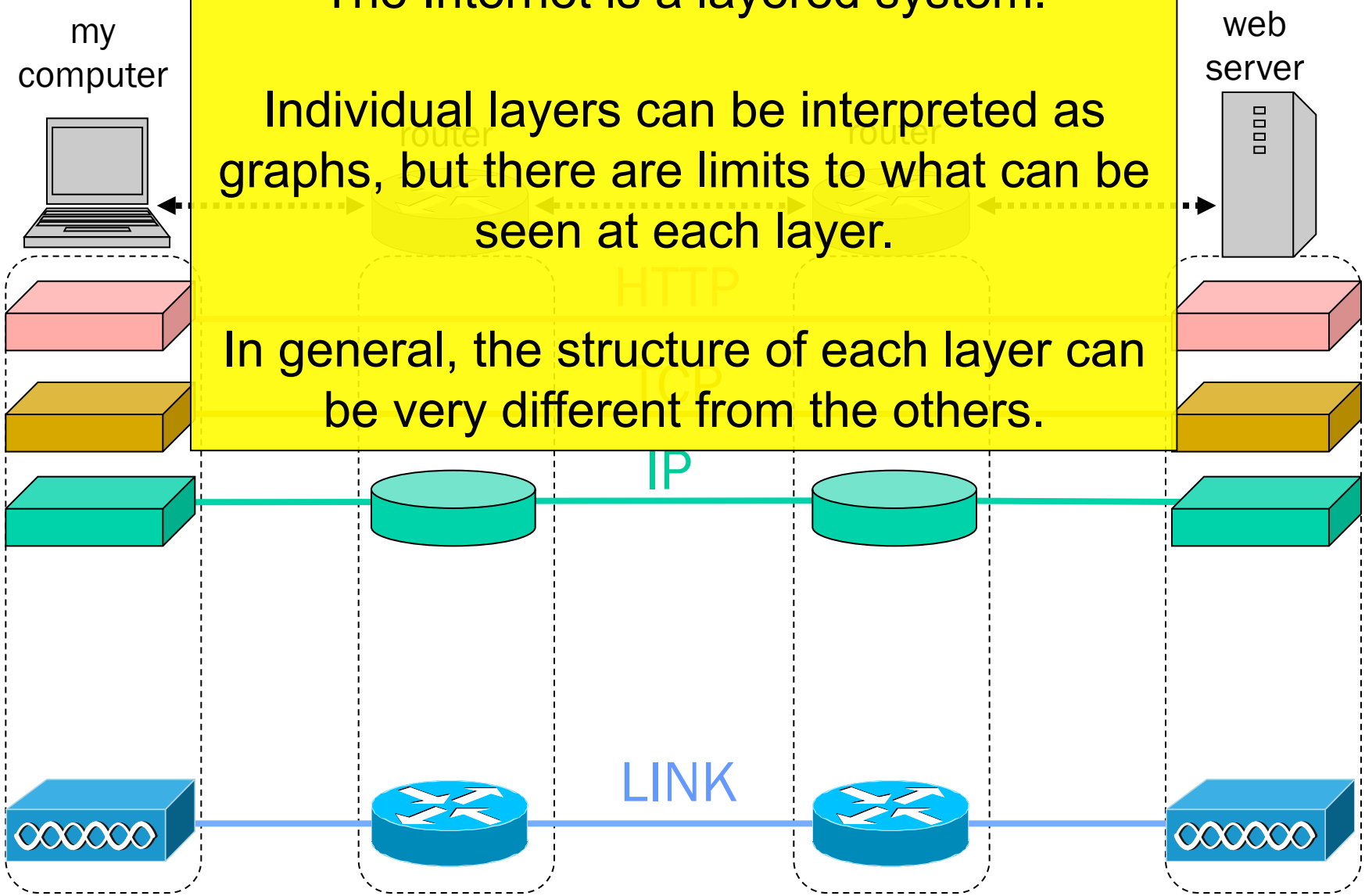


The Internet is NOT a graph!

The Internet is a layered system.

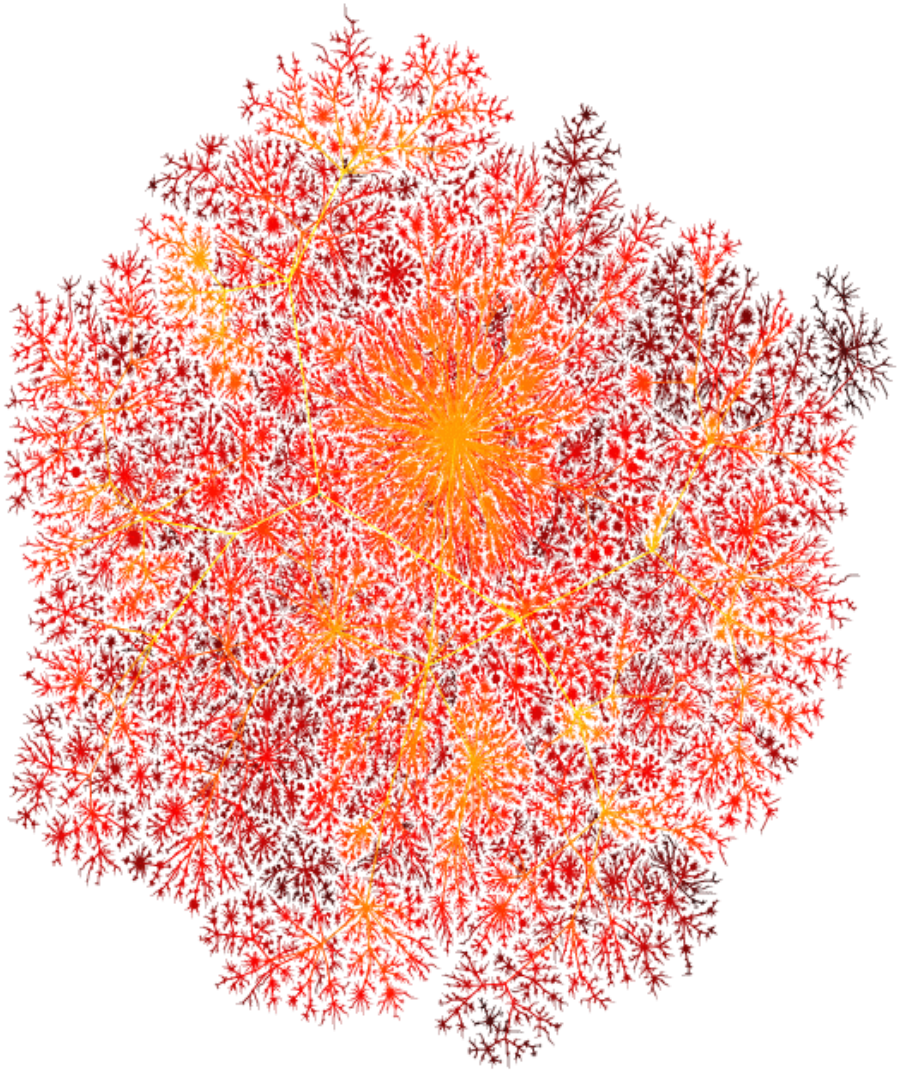
Individual layers can be interpreted as graphs, but there are limits to what can be seen at each layer.

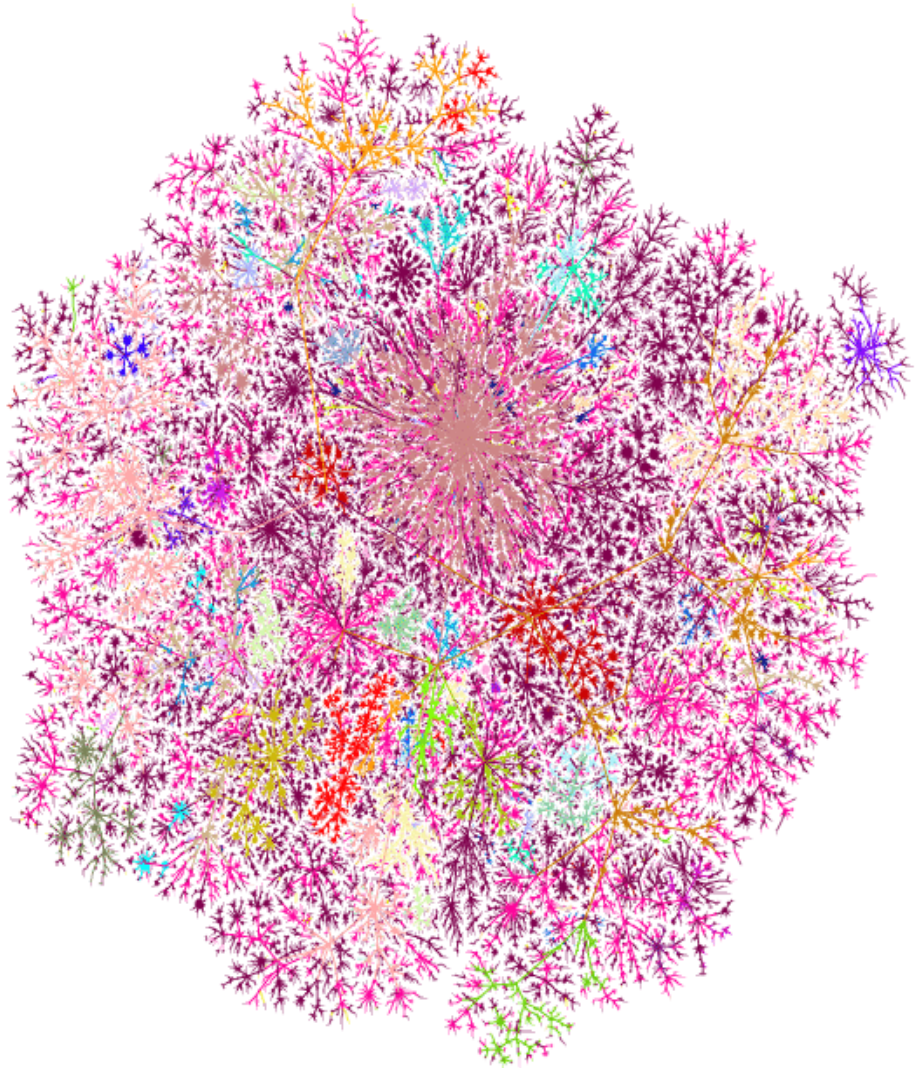
In general, the structure of each layer can be very different from the others.



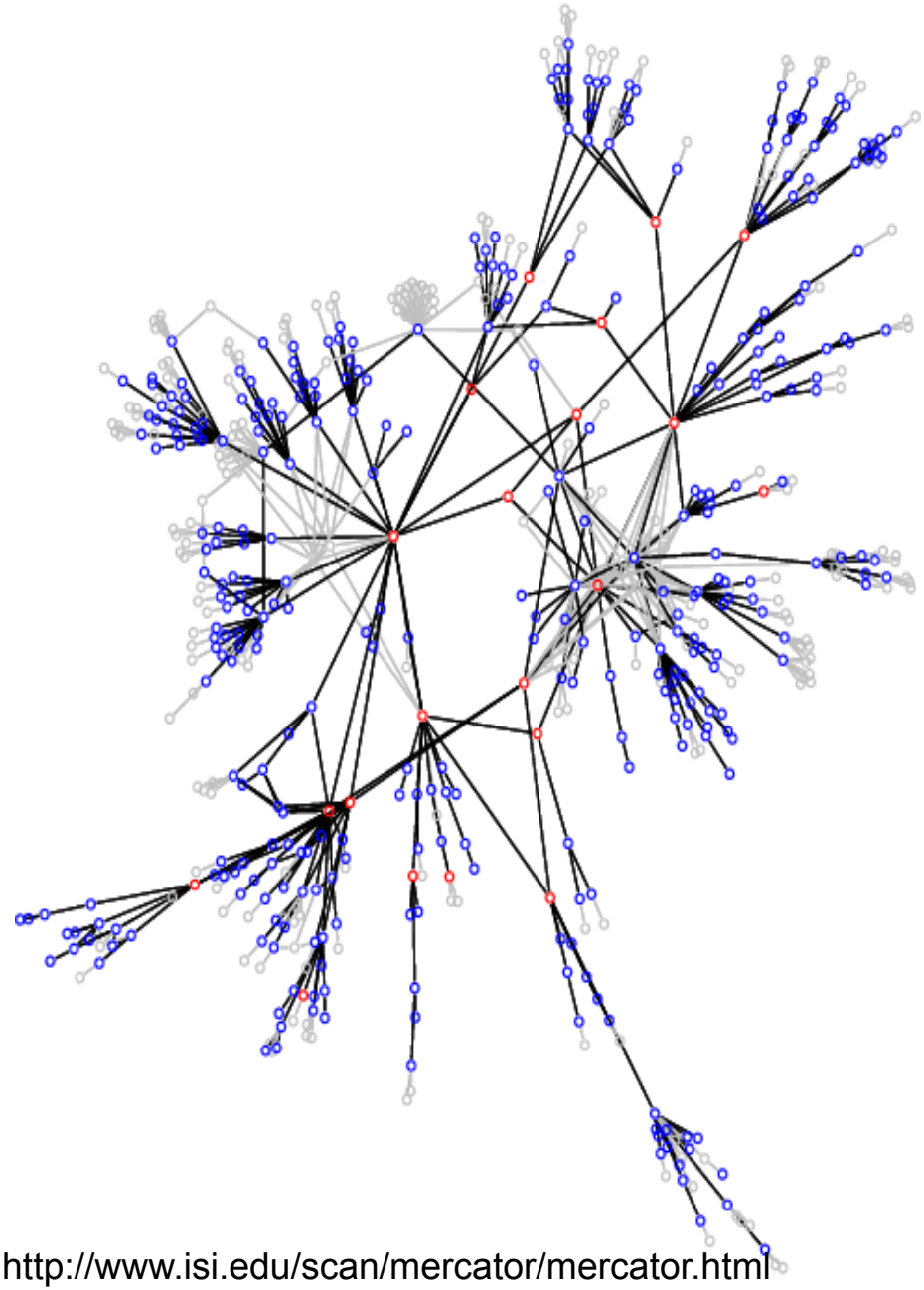
Unfortunately, direct inspection of Internet topology is generally NOT possible

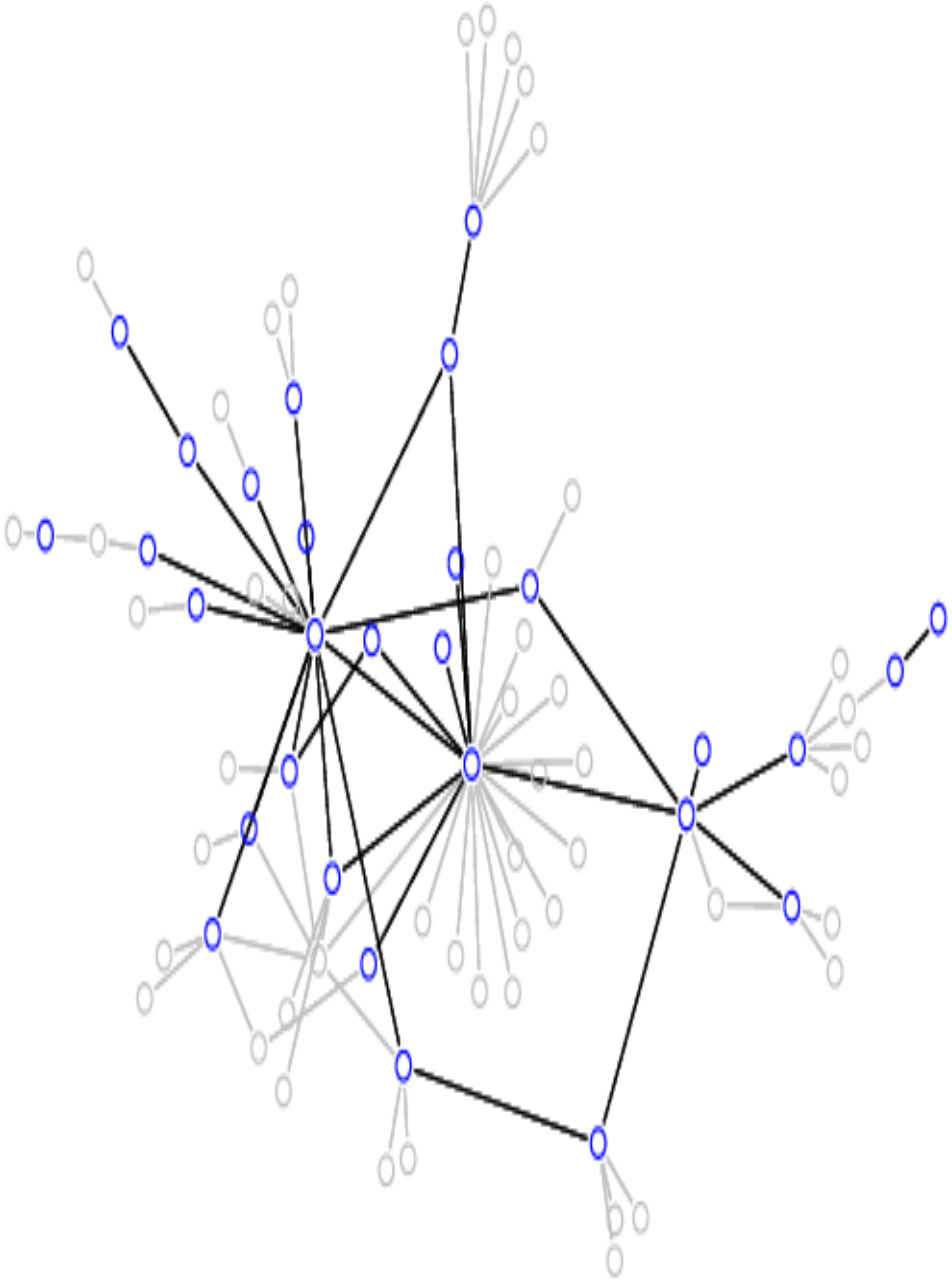
- **No single vantage point from which to observe everything**
- **Economic incentive for ISPs to obscure network structure**
- **Recent trends**
  - **Empirical measurement studies**
  - **Generative models that "fit" measured statistics**
- example: **`traceroute`** tool
  - Discovers compliant (i.e., IP) routers along path between selected network host computers
  - Originally designed to record the end-to-end traffic path from source to destination



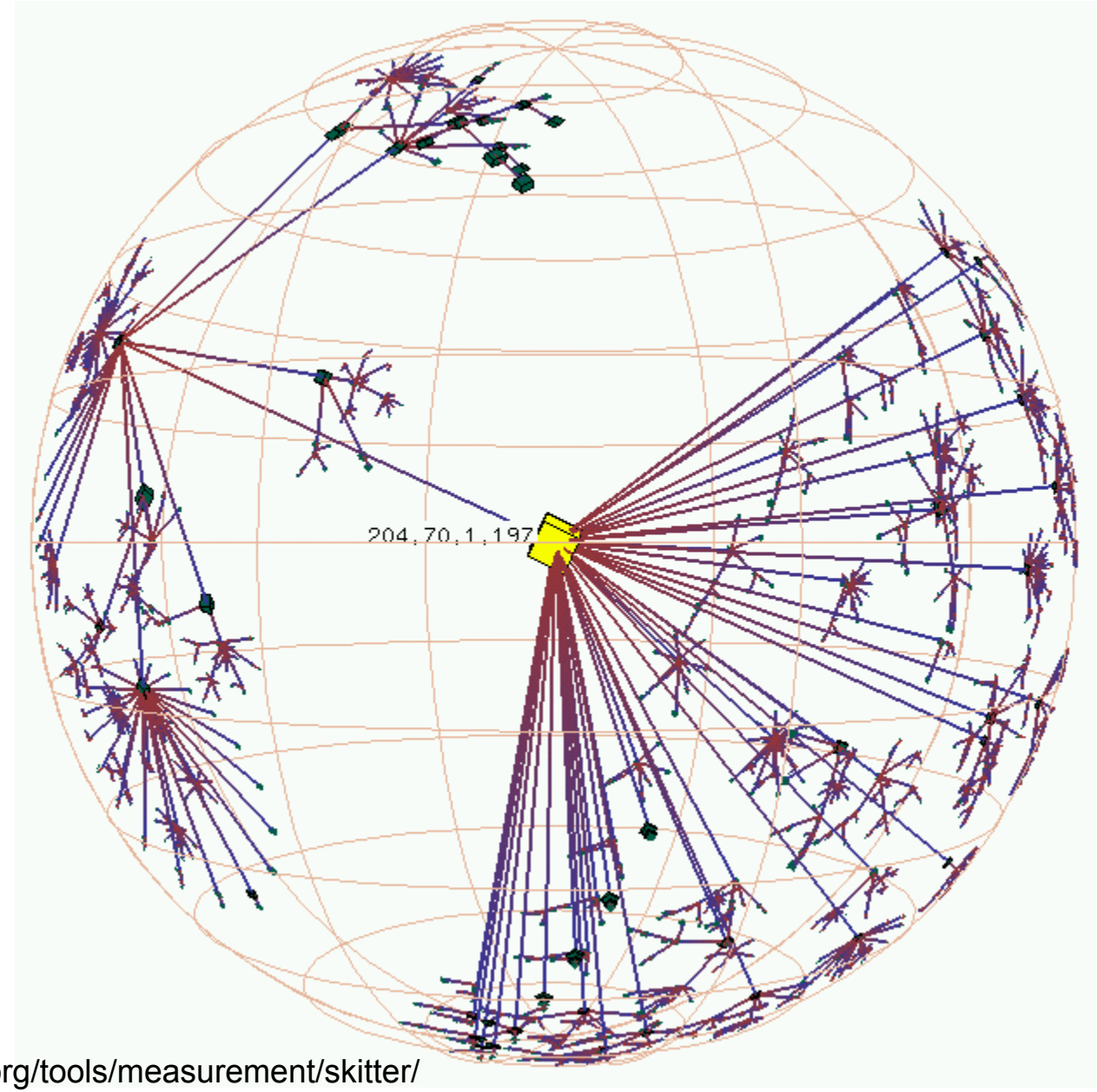


http://research.lumeta.com/ches/map/

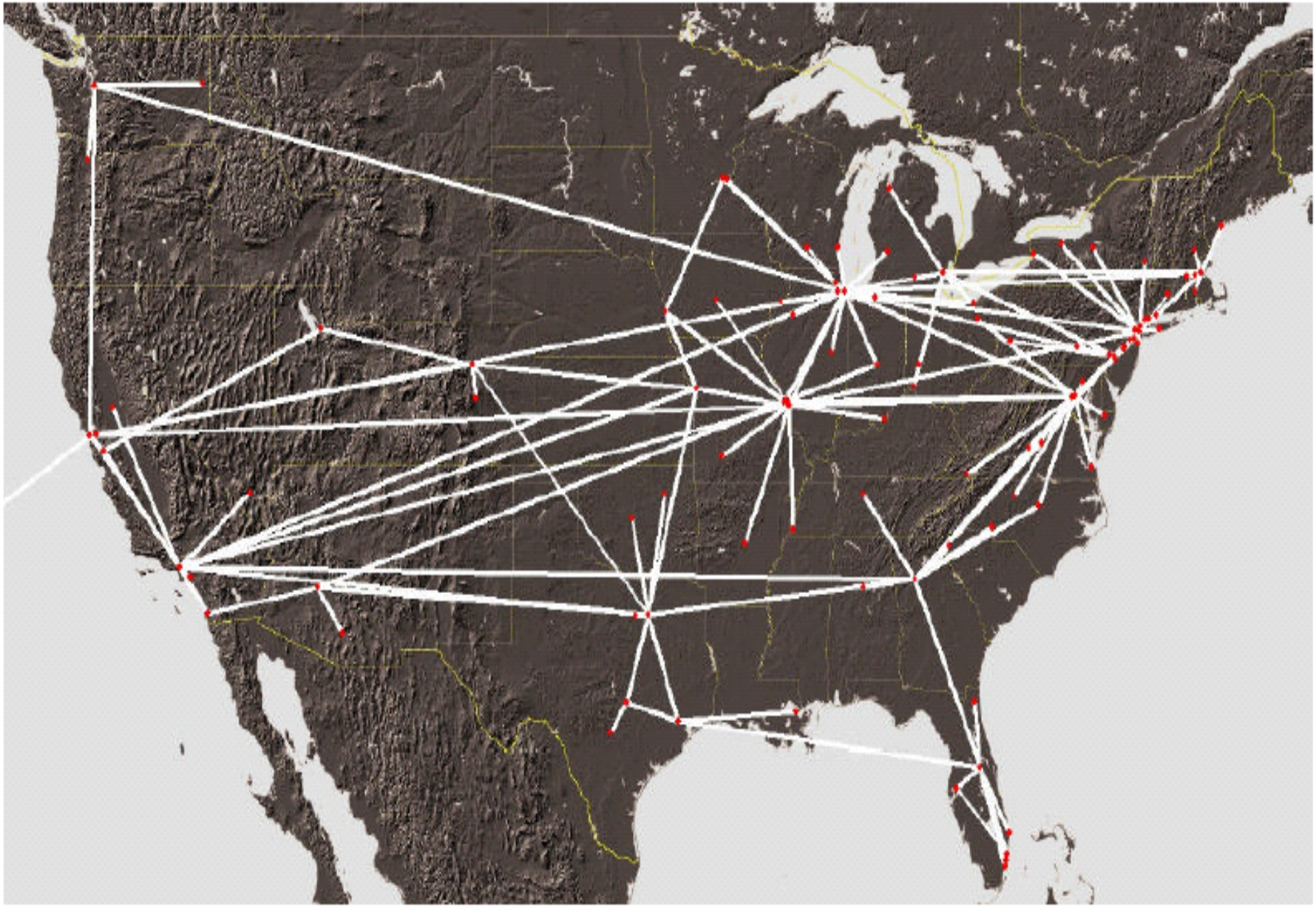




http://www.isi.edu/scan/mercator/mercator.html



http://www.caida.org/tools/measurement/skitter/



Background image courtesy JHU, applied physics labs

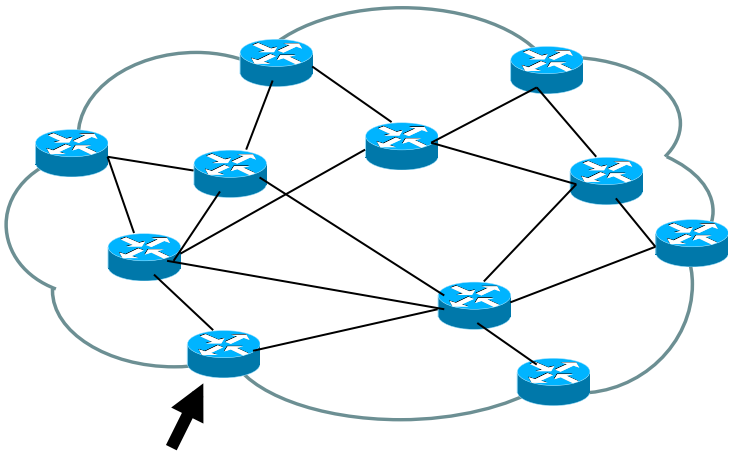
http://www.cs.washington.edu/research/networking/rocketfuel/bb

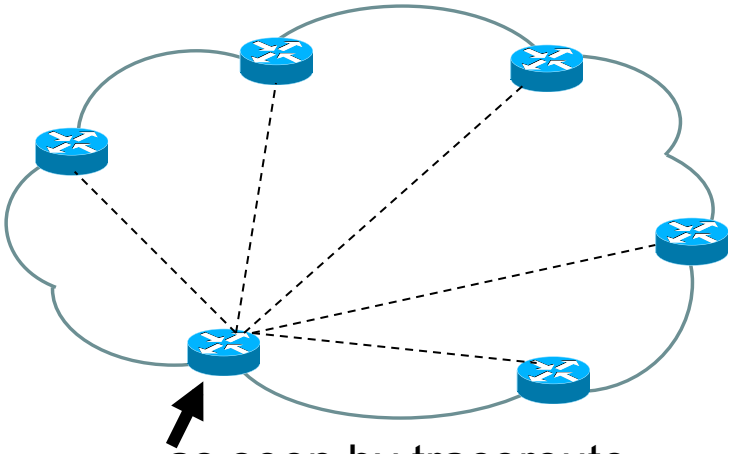
These are all nice graphs, but do we believe them?

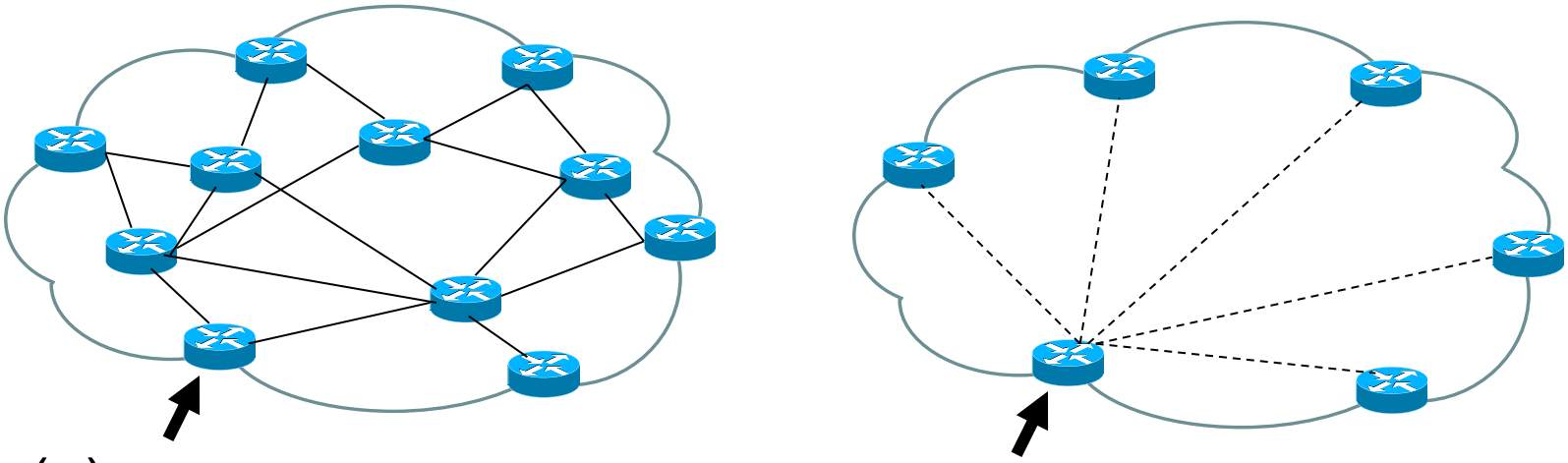
- How was the graph data collected?
- Are there biases, systematic errors, or idiosyncrasies, in the measurement data?
- What types of analyses can the graph data support?
- How does uncertainty in the underlying graph data affect the analyses that you perform?

Using `traceroute` to infer topology

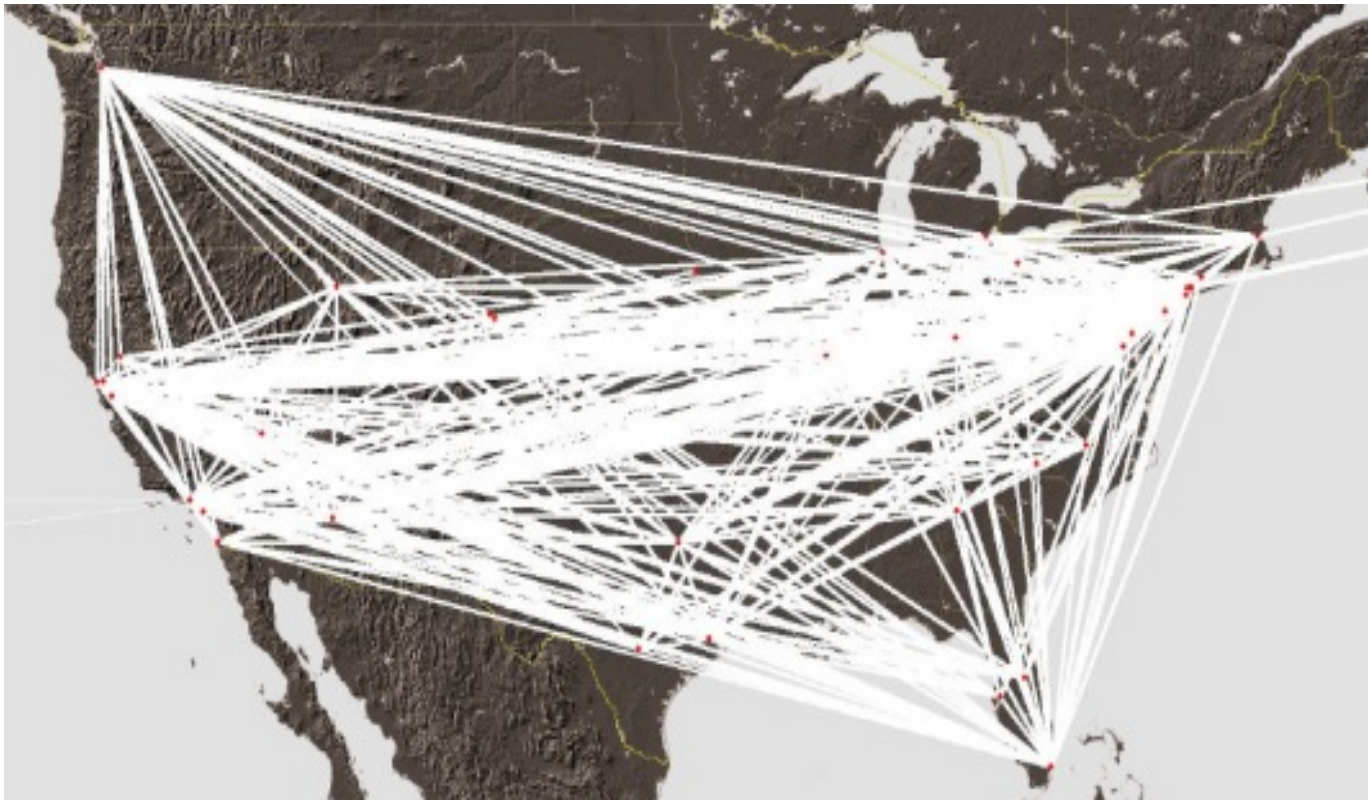
- traceroute-based measurements are ambiguous
  - traceroute is strictly about IP-level connectivity
  - traceroute cannot distinguish between high connectivity nodes that are for real and that are fake and due to underlying Layer 2 (e.g., Ethernet, ATM) or Layer 2.5 technologies (e.g., MPLS)



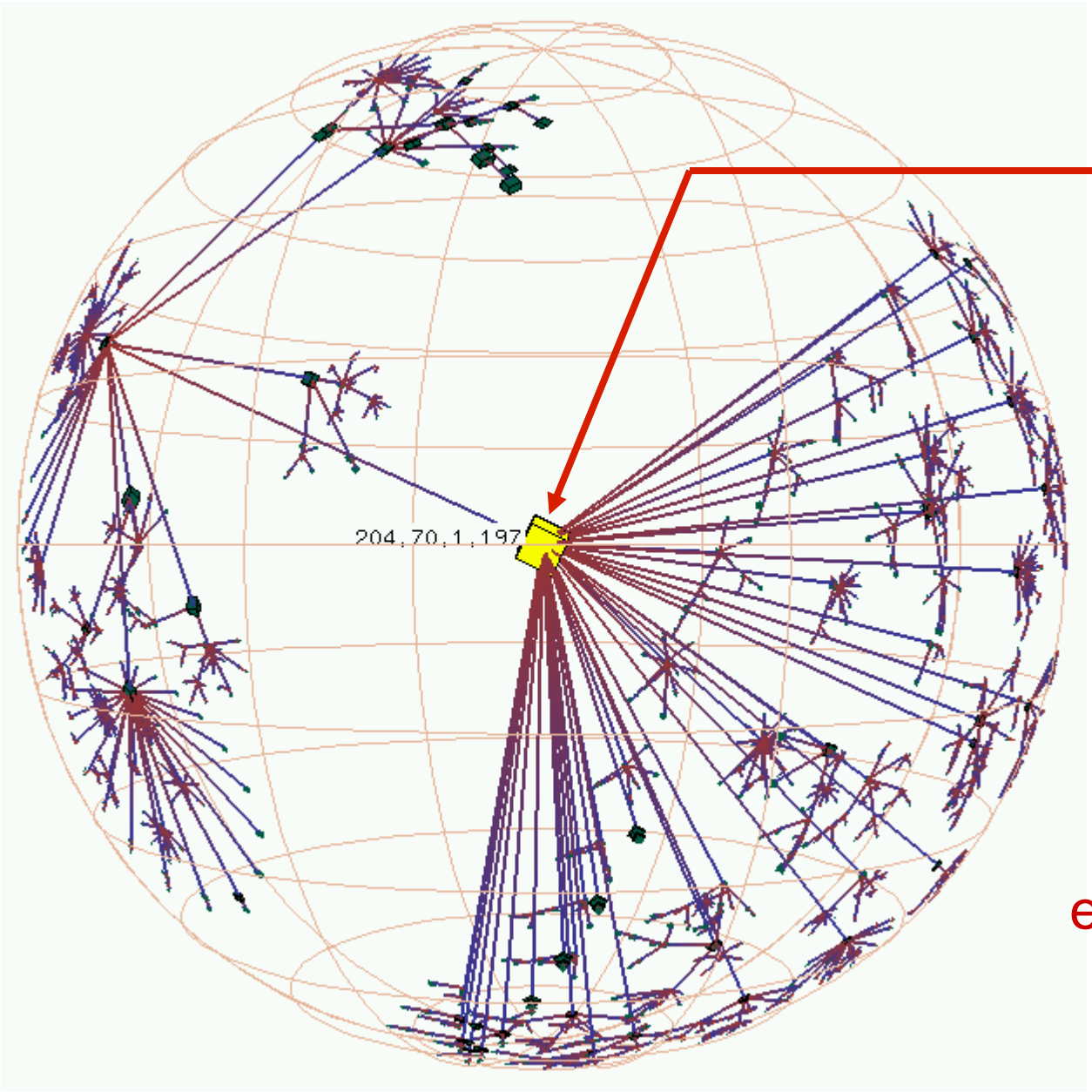




(a)



(b)



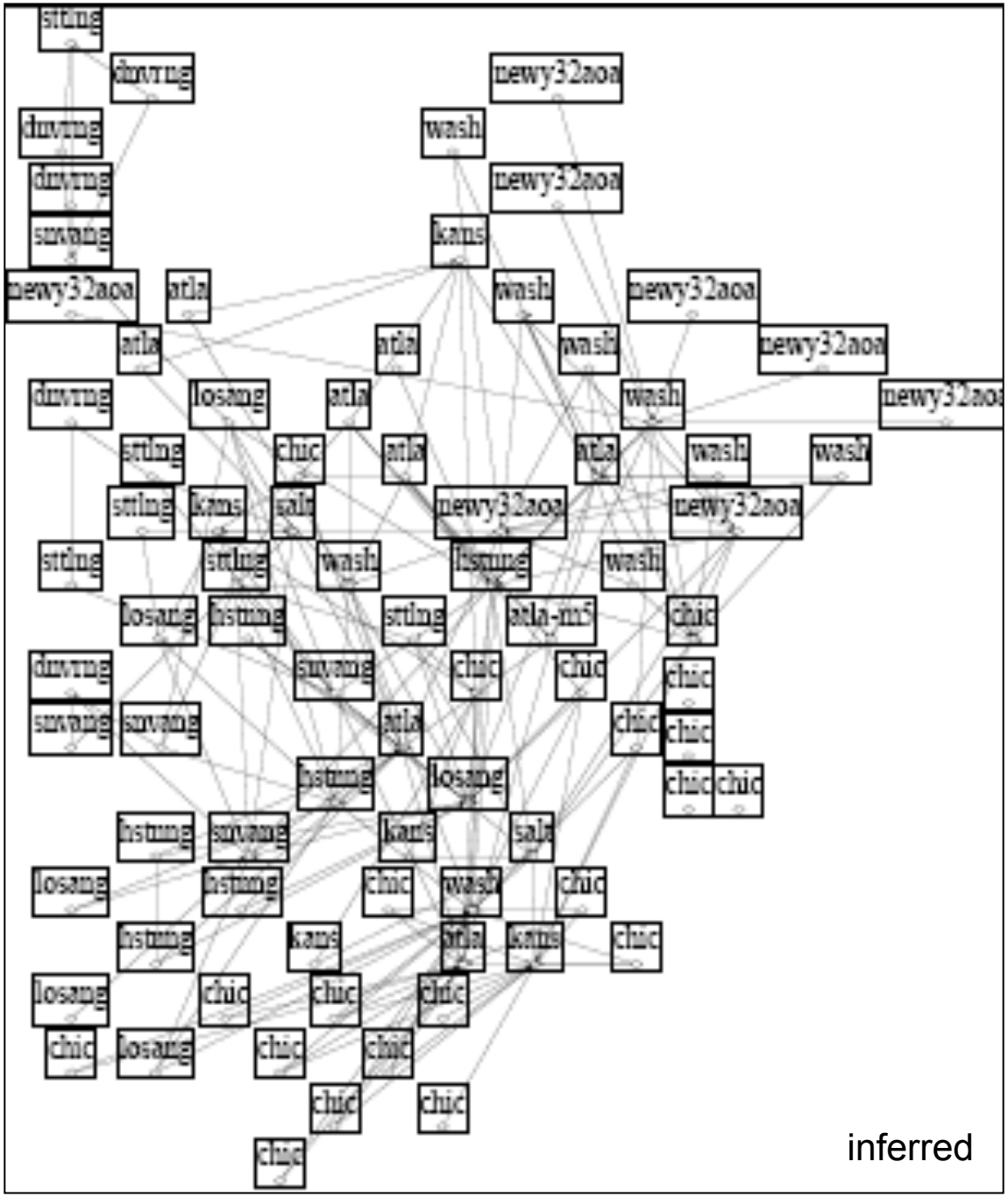
- www.savvis.net
- managed IP and hosting company
- founded 1995
- offering “private IP with ATM at core”

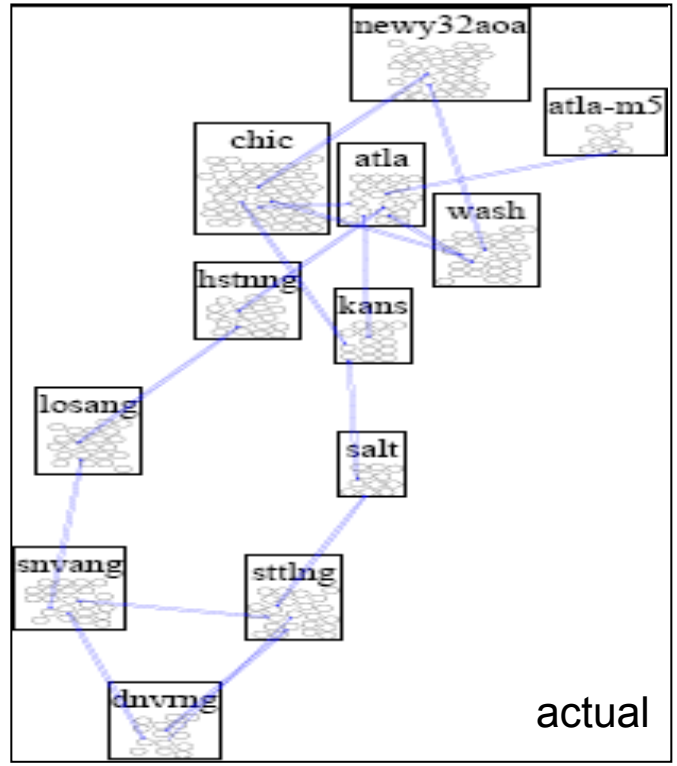
Possible that this “node” is an entire network! (not just a router)

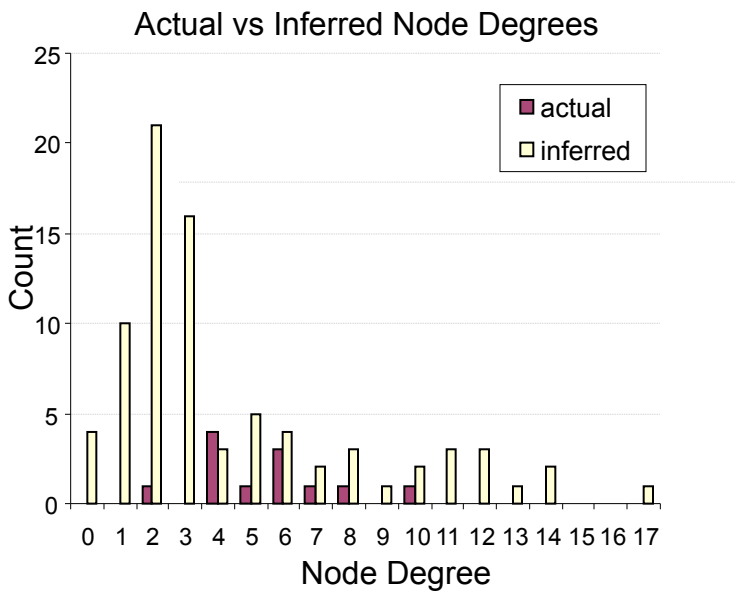
http://www.caida.org/tools/measurement/skitter/

Using `traceroute` to infer topology

- traceroute-based measurements are inaccurate
  - Requires some guesswork in deciding which IP addresses/interface cards refer to the same router ("alias resolution" problem)







Source: R. Sherwood, A. Bender, and N. Spring, DisCarte: A disjunctive Internet cartographer, *Proc. ACM SIGCOMM*, 2008.

Using `traceroute` to infer topology

- traceroute-based measurements are incomplete
  - traceroute experiments unable to detect high degree nodes at edge, if they exist
  - IP-level connectivity is more easily/accurately inferred the closer the routers are to the traceroute source(s)
    - notion of bias, most studied (but least important?)
    - node degree distribution is inferred to be power-law even when the actual distribution is not

Using `traceroute` to infer topology

- traceroute-based measurements are ambiguous
  - false positives on high degree nodes (due to underlying Layer 2 technologies)
- traceroute-based measurements are inaccurate
  - systematic error in estimating node degree (aliasing)
- traceroute-based measurements are incomplete
  - false negatives on high degree nodes at edge

Conclusion: node degree data inferred from traceroute measurements is insufficient to support statistical modeling

Conclusion: models that emphasize replicating router-level node degree statistics are unjustified

Reference: W. Willinger, D. Alderson, and J.C. Doyle. "Mathematics and the Internet: A source of enormous confusion and great potential," Notices of the American Mathematical Society 56(5):286-299, May 2009.

“Inferring Relevant Social Networks from Interpersonal Communication”
M. De Choudhury, W. A. Mason, J. M. Hofman, D. J. Watts, 2010,
WWW 2010, April 26–30, 2010, Raleigh, North Carolina, USA.

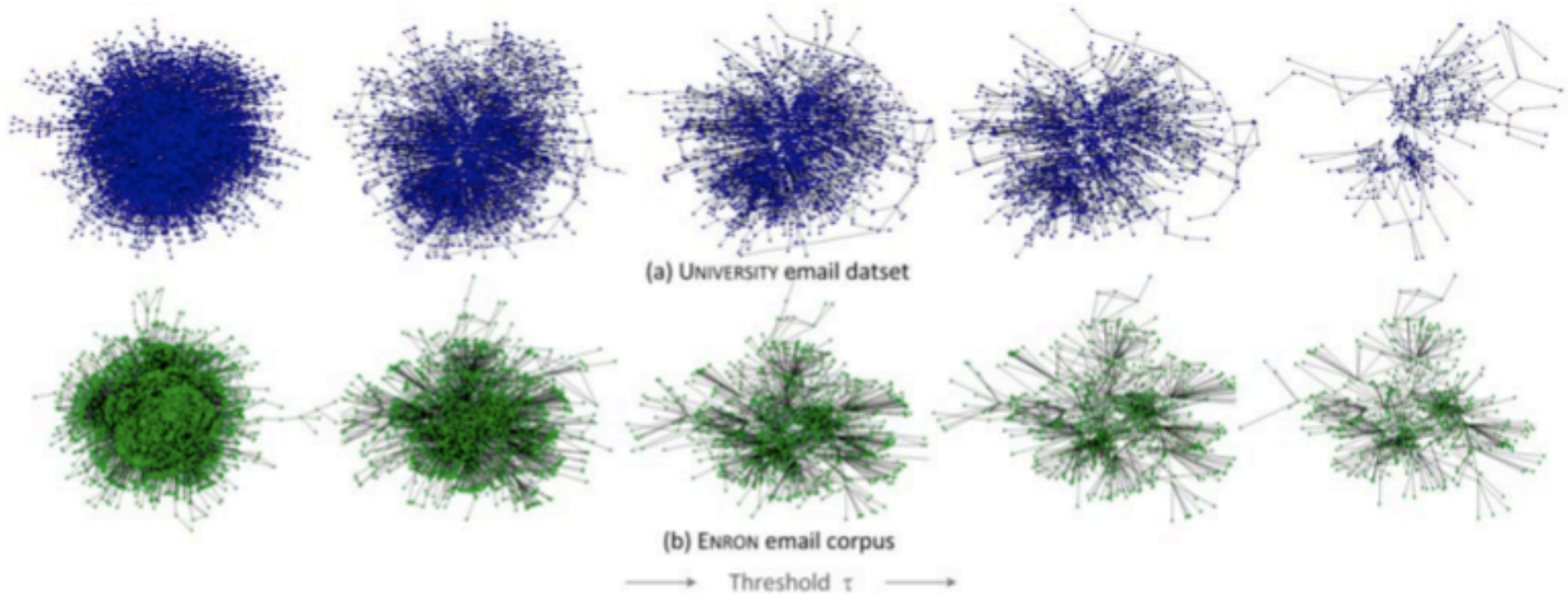
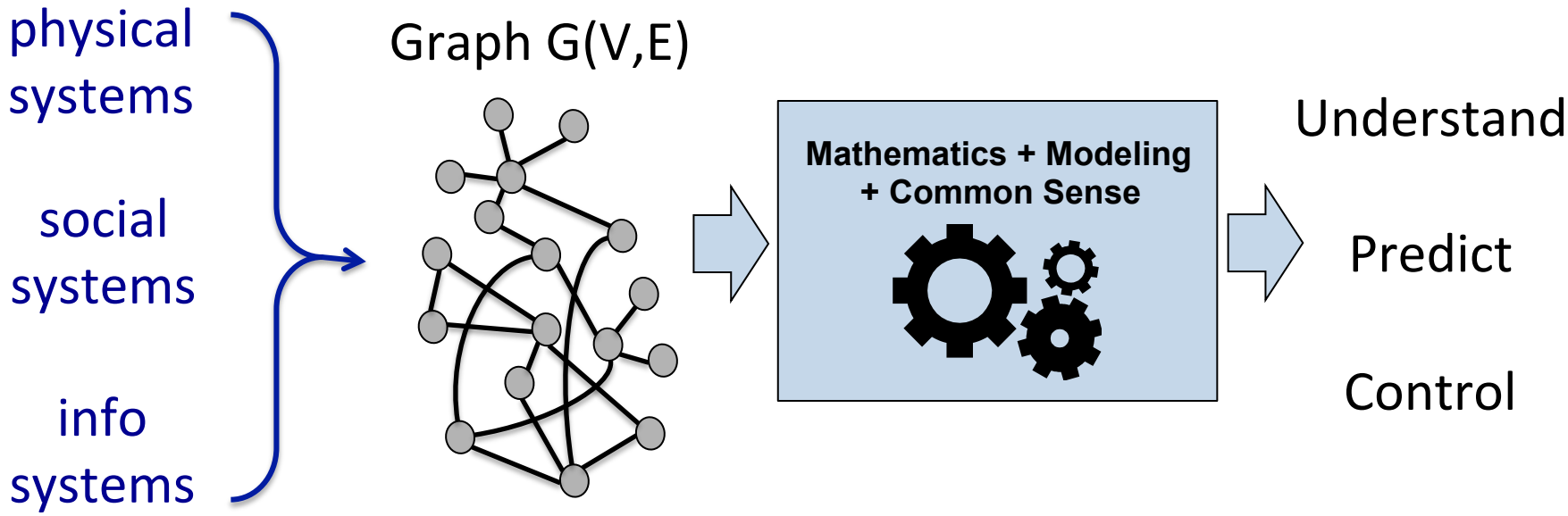


Figure 1: Topology of the largest components over various choices of threshold conditions for (a) a dataset based on email server logs at a US university, and (b) the Enron email corpus. Significant changes in topology are observed as the thresholding condition of the network is varied.

“different choices of the threshold correspond to dramatically different network structures”

“there is no one ‘true’ social network”

Pitfall #2: Does the graph capture the essence of the system?



- Many systems of interest are not arbitrary in their structure, but have evolved in order to achieve some *purpose* or *function*
- Ignoring this function can yield misleading results
- System function provides a natural lens for validating a graph representation of a system (i.e., could it work?)

The New York Times Reprints

March 20, 2010

http://www.nytimes.com/2010/03/20/world/asia/21grid.html

Academic Paper in China Sets Off Alarms in U.S.

By JOHN MARKOFF and DAVID BARBOZA

It came as a surprise this month to Wang Jianwei, a graduate engineering student in Liaoning, China, that he had been described as a potential cyberwarrior before the United States Congress.

Larry M. Wortzel, a military strategist and China specialist, told the House Foreign Affairs

Cascade-based attack vulnerability on the US power grid

Jian-Wei Wang, and Li-Li Rong

Institute of System Engineering, Dalian University of Technology, 2 Ling Gong Rd., Dalian 116024, Liaoning, PR China

Abstract

The vulnerability of real-life networks subject to intentional attacks has been one of the outstanding challenges in the study of the network safety. Applying the real data of the US power grid, we compare the effects of two different attacks for the network robustness against cascading failures, i.e., removal by either the descending or ascending orders of the loads. Adopting the initial load of a node j to be $L_j=[k_j(\Sigma_m \Gamma_{jkm})]^\alpha$ with k_j and Γ_j being the degree of the node j and the set of its neighboring nodes, respectively, where α is a tunable parameter and governs the strength of the initial load of a node, we investigate the response of the US power grid under two attacks during the cascading propagation. In the case of $\alpha<0.7$, our investigation by the numerical simulations leads to a counterintuitive finding on the US power grid that the attack on the nodes with the lowest loads is more harmful than the attack on the ones with the highest loads. In addition, the almost same effect of two attacks in the case of $\alpha=0.7$ may be useful in furthering studies on the control and defense of cascading failures in the US power grid.

Keywords: Cascading failure; Attack; US power grid; Critical threshold; Tunable parameter

Safety Science **47** (2009) 1332-1336.

J-W Wang, and L-L Rong (2009) “Cascade-based attack vulnerability on the US power grid,” *Safety Science* **47,** 1332-1336.

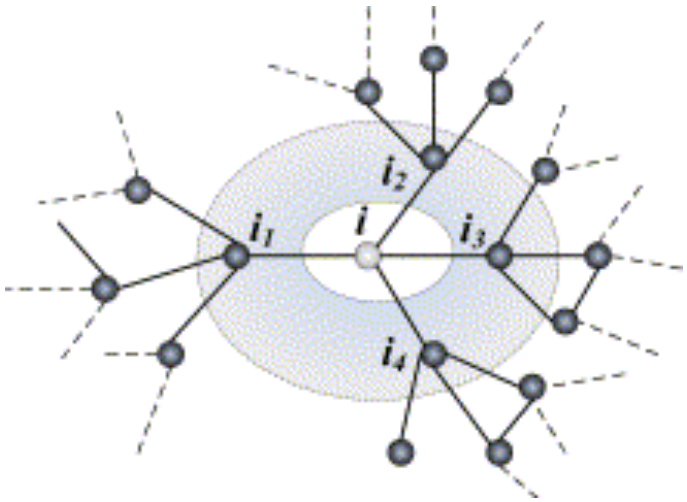


Fig. 1. The scheme illustrates the correlation between the initial load of a node *i* and its neighboring nodes, i.e., nodes i_1, i_2, i_3, and i_4.

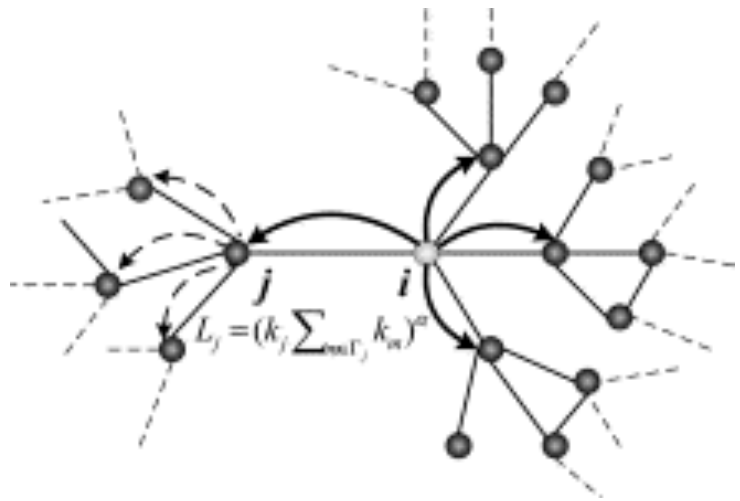


Fig. 2. The scheme illustrates the load redistribution triggered by an node-based attack. Node *i* is removed and the load on it is redistributed to the neighboring nodes connecting to node *i*. Among these neighboring nodes, the one with the higher load will receive the higher shared load from the broken node.

$$L_j = \left[k_j \left(\sum_{m \in \Gamma_j} k_m \right) \right]^\alpha$$

$$k_j = \deg \textit{ of node } j$$

$$\Gamma_j = \textit{neighbors of node } j$$

$$\alpha = \textit{tunable parameter}$$

$$\Delta L_{ji} = L_i \frac{L_j}{\sum_{m \in \Gamma_j} L_m}$$

1. graph connectivity only
2. borrowed data, at face value
3. inconsistent with engineering reality

Modeling Electric Power Grid Interdiction

- System function depends on all *generators*, *high-voltage lines* and *transformers*, *buses* and *substations* in a region
- Behavior comes from solving a set of *Power Flow Models* to *minimize unserved demand* for energy
  - Kirchhoff's current law
  - Line thermal capacities; Generating capacities
  - Admittance constraints
  - Generating cost per unit; Shedding cost per customer
  - Demand variation and Component Restoration over time
- Can look for vulnerabilities by taking the perspective of an *Attacker* who uses limited resources to attack components (e.g., # of substations or lines)

- Salmerón, J., Wood, K. and Baldick, R., 2009, "Worst-Case Interdiction Analysis of Large-Scale Electric Power Grids," *IEEE Transactions on Power Systems*, **24**, pp. 96-104.
- Salmeron, J., Wood, K. and Baldick, R., 2004, "Analysis of Electric Grid Security Under Terrorist Threat," *IEEE Transactions on Power Systems*, 19, pp. 905-912.

Modeling Electric Power Grid Interdiction

Operator's Model: **an industry-standard optimal power-flow model (OPF, DC approximation) to find the best solution, even in the presence of component loss (e.g., accidents or Mother Nature).**

$$\min_{P^{Gen},P^{Line},S,\theta} \sum_i \sum_{g\in G_i} h(P_g^{Gen}) + \sum_i \sum_c f(S_{ic}) \quad \text{subject to:}$$

$$\begin{cases} P_l^{Line} = B_l(\theta_{o(l)} - \theta_{d(l)}), & \forall l \in L \\ \sum_{g\in G_i} P_{i,g}^{Gen} - \sum_{l|o(l)=i} P_l^{Line} + \sum_{l|d(l)=i} P_l^{Line} = \sum_c (d_{ic} - S_{ic}), & \forall i \\ -\overline{P}_l^{Line} \le P_l^{Line} \le \overline{P}_l^{Line}, & \forall l \in L \\ \underline{P}_{i,g}^{Gen} \le P_{i,g}^{Gen} \le \overline{P}_{i,g}^{Gen}, & \forall i, \forall g \in G_i \\ 0 \le S_{ic} \le d_{ic}, & \forall i, c \end{cases}$$

i: bus, l: line, g: generator, c: customer sector

P^{Line}, P^{Gen}: power (MW) S: power shed θ : bus phase

Rest of this talk:

What is most {important, vital, critical} in the network?

↓

Where are the vulnerabilities?

What are the highest risks?

↙ ↘

Where to target?

What to protect?

For each of these, how to prioritize investment?

(e.g., use of lists to rank priorities)

Objective: infrastructure *resilience*

National Strategy for Homeland Security (2007)

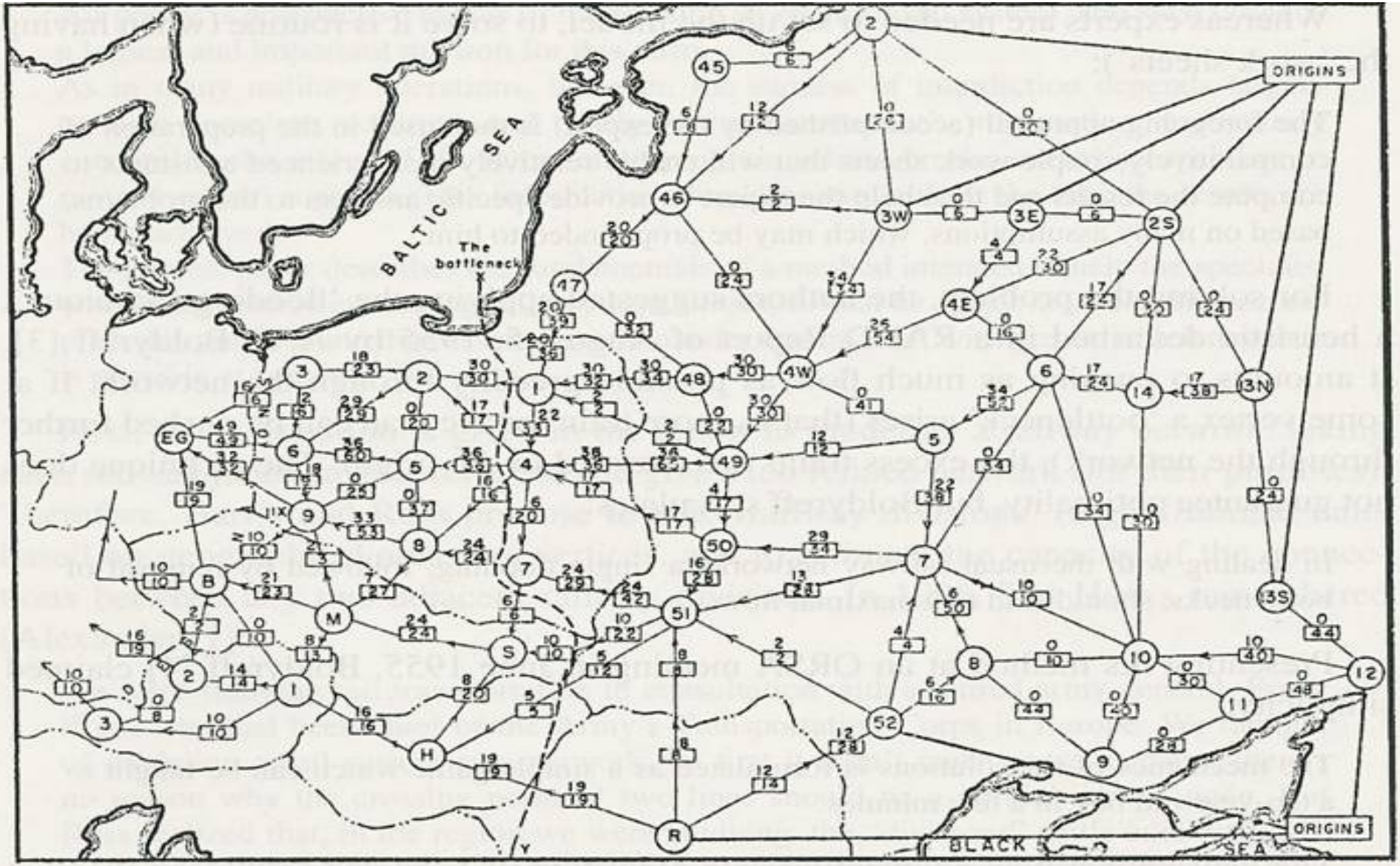
"We will not be able to deter all terrorist threats, and it is impossible to deter or prevent natural catastrophes. We can, however, mitigate the Nation's vulnerability to acts of terrorism, other man-made threats, and natural disasters by **ensuring the structural and operational resilience** of our critical infrastructure and key resources" (p. 27)

"We must now focus on the **resilience of the system as a whole** – an approach that centers on investments that make the system better able to absorb the impact of an event without losing the capacity to function" (p.28)

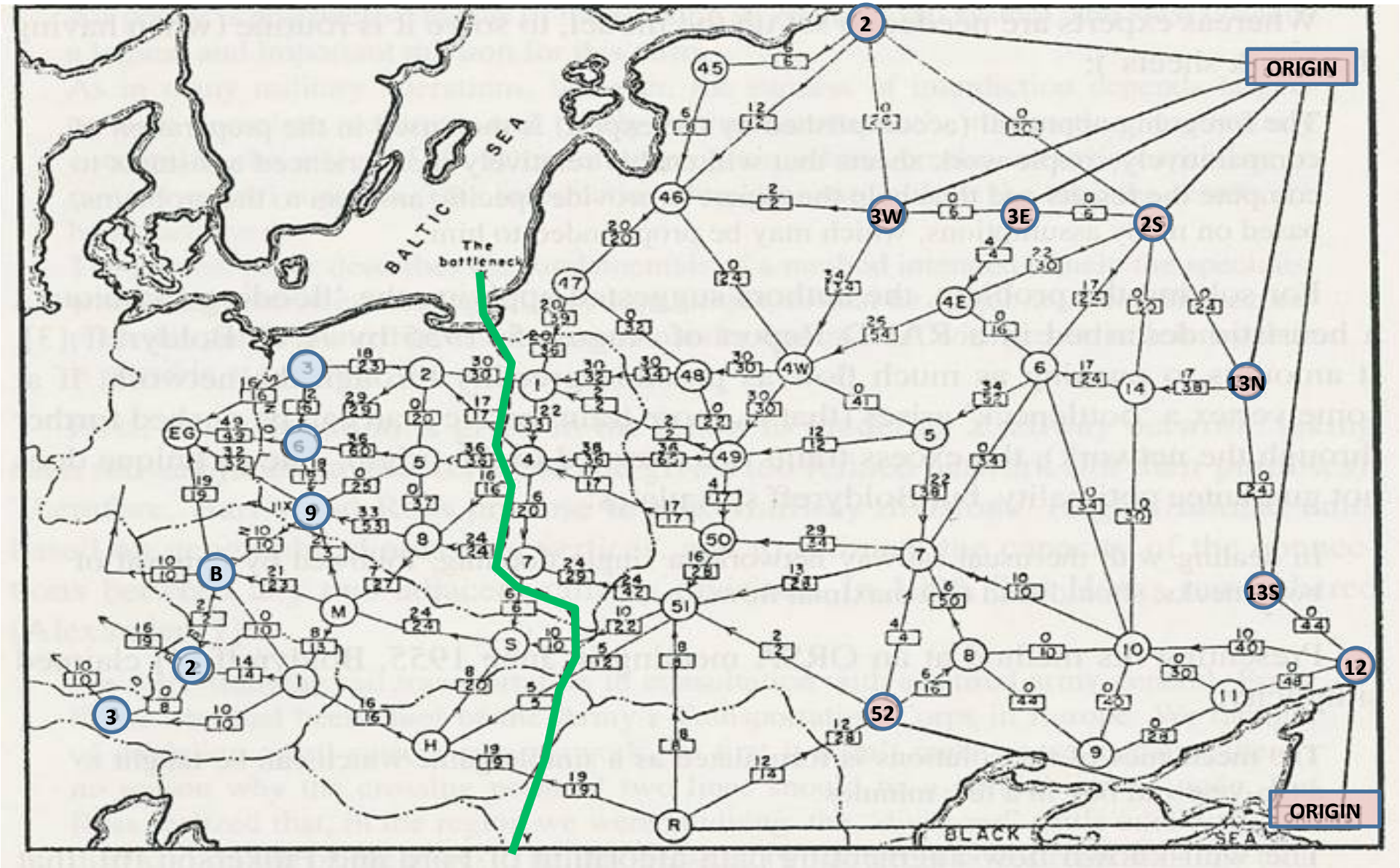
The Russian Rail Network (circa 1955)

Data from Figure 7 of:

Harris, T.E., and Ross, F.S. (1955), *Fundamentals of a Method for Evaluating Rail Net Capacities* (SECRET, declassified 1999), RM-1573, RAND Corp.


ORIGINS
BALTIC SEA
The bottleneck
BLACK SEA
ORIGINS

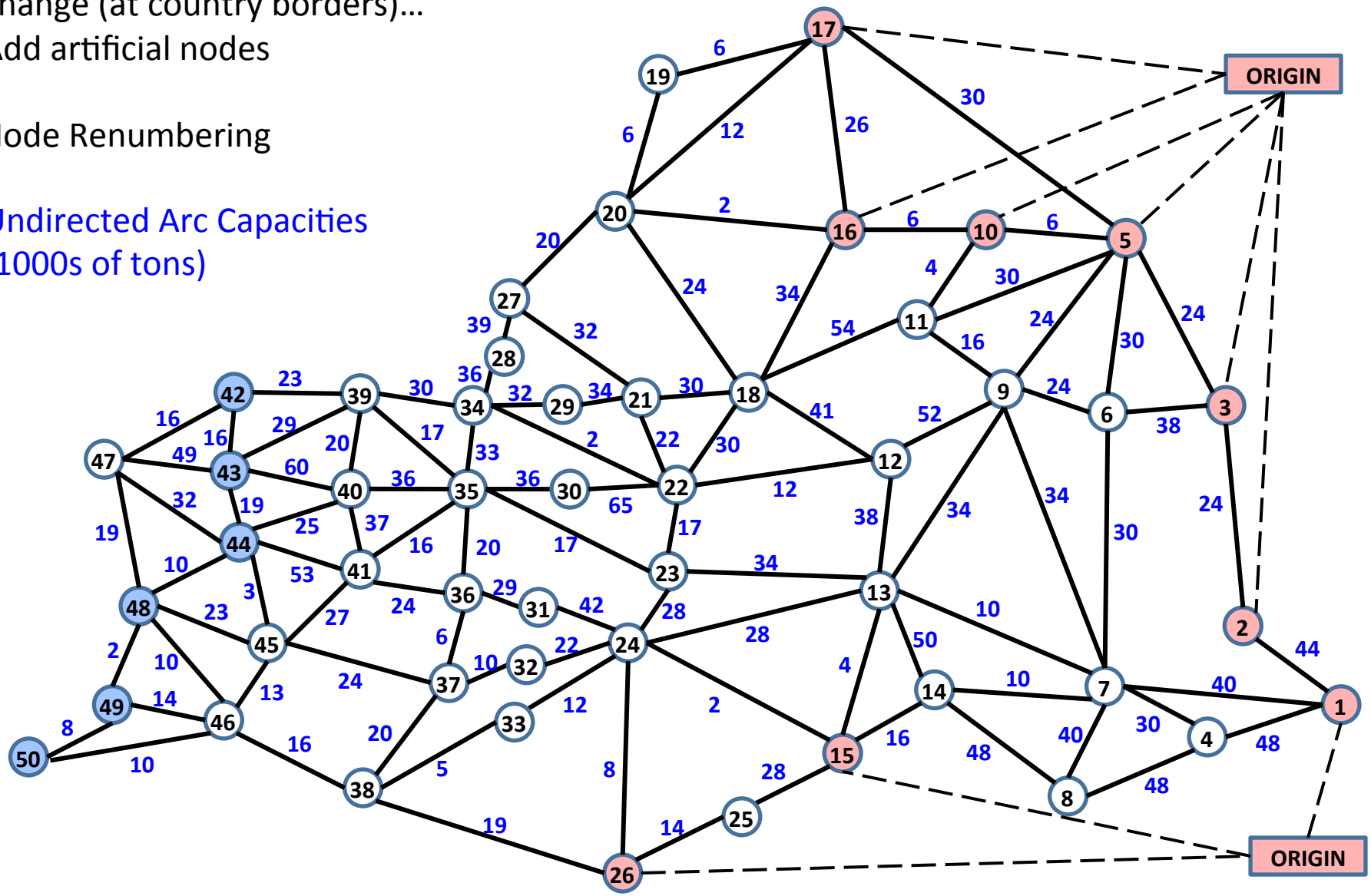
destination nodes minimum capacity cut origin nodes



7 places where arc capacities change (at country borders)…
Add artificial nodes

Node Renumbering

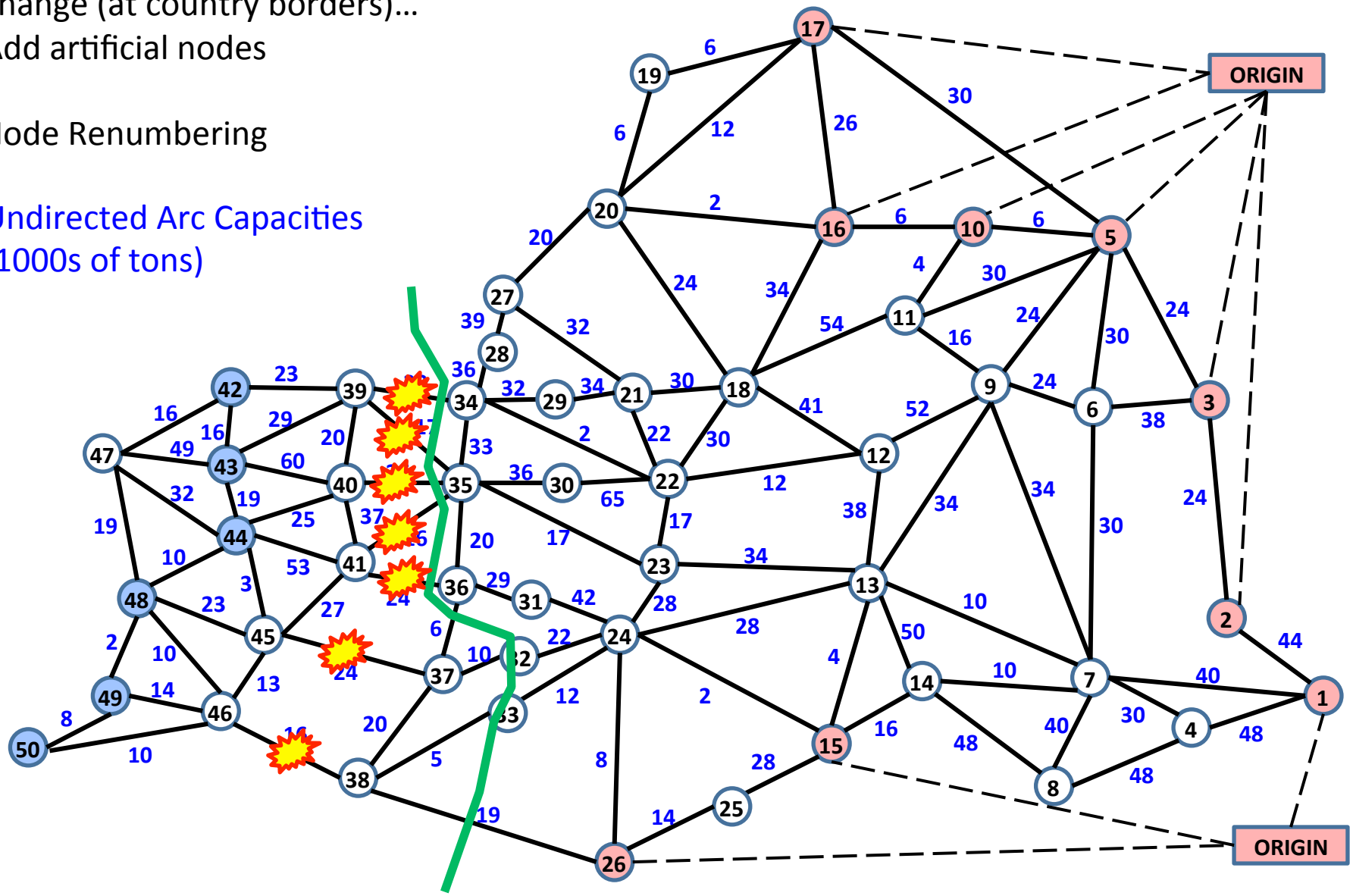
Undirected Arc Capacities (1000s of tons)



7 places where arc capacities change (at country borders)…
Add artificial nodes

Node Renumbering

Undirected Arc Capacities (1000s of tons)



Working Definition

A *most-vital arc in a maximum-flow problem* is one whose removal decreases the resulting maximum flow by the greatest amount.

Heuristics for determining what is most vital

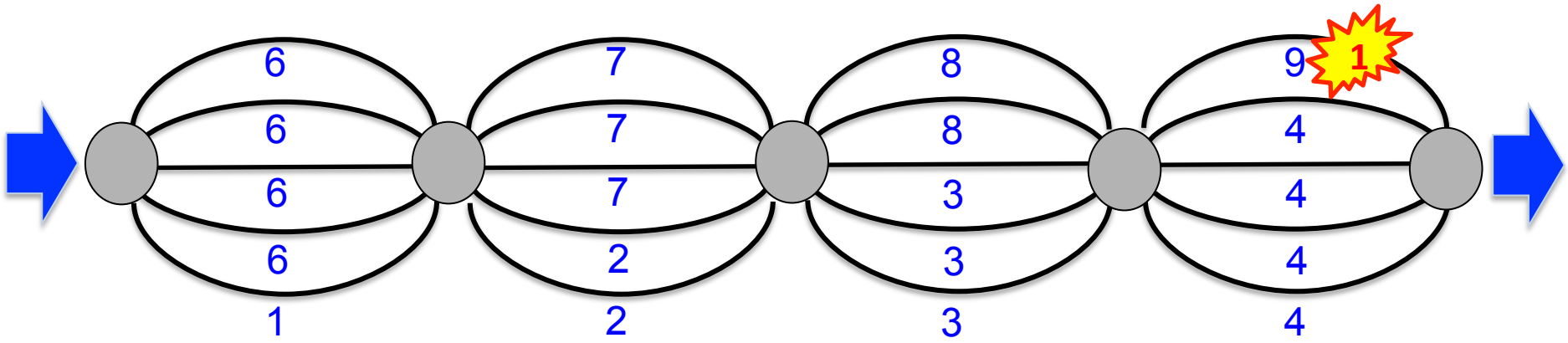
(Ahuja, Magnanti, and Orlin, "Network Flows", Prentice-Hall, 1993)

- An arc having the *largest capacity*
- An arc carrying the *largest flow in the optimal solution*
- An arc having the *largest capacity in a minimum-capacity cut*
- *Any most-vital arc is in some minimum-capacity cut*

In general, none of these heuristics work!

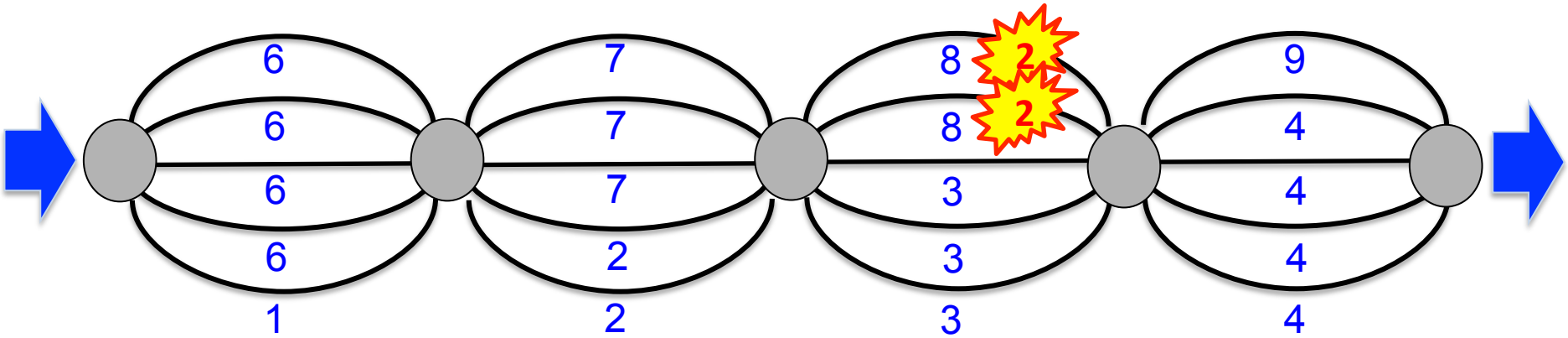
Determining a most vital arc requires solving a *maximum-flow interdiction problem*, which is an ILP problem of size comparable to the original maximum-flow problem (Wood 1993).

Why simple priority lists don't work



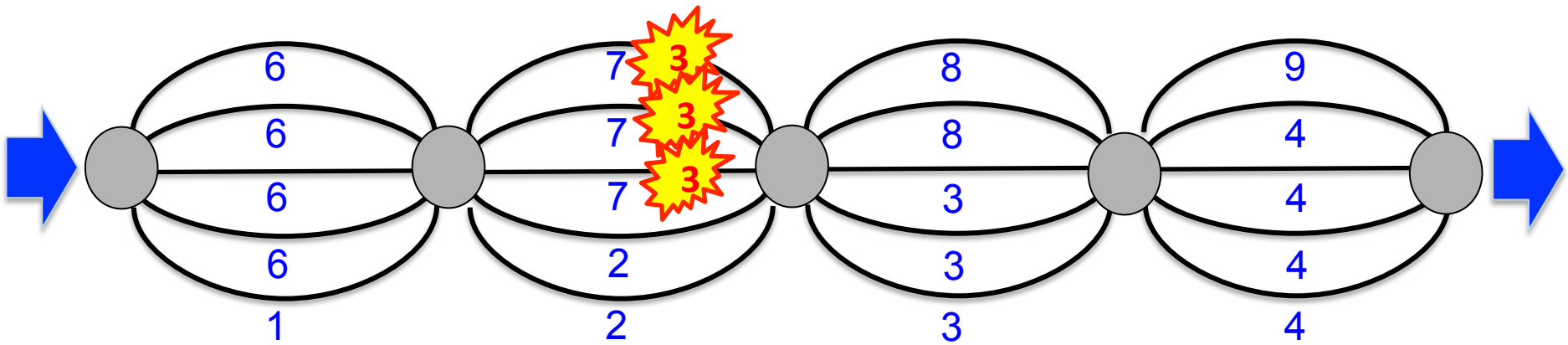
- With 0 attacks, capacity = 25
- With 1 attack, capacity = 16

Why simple priority lists don't work



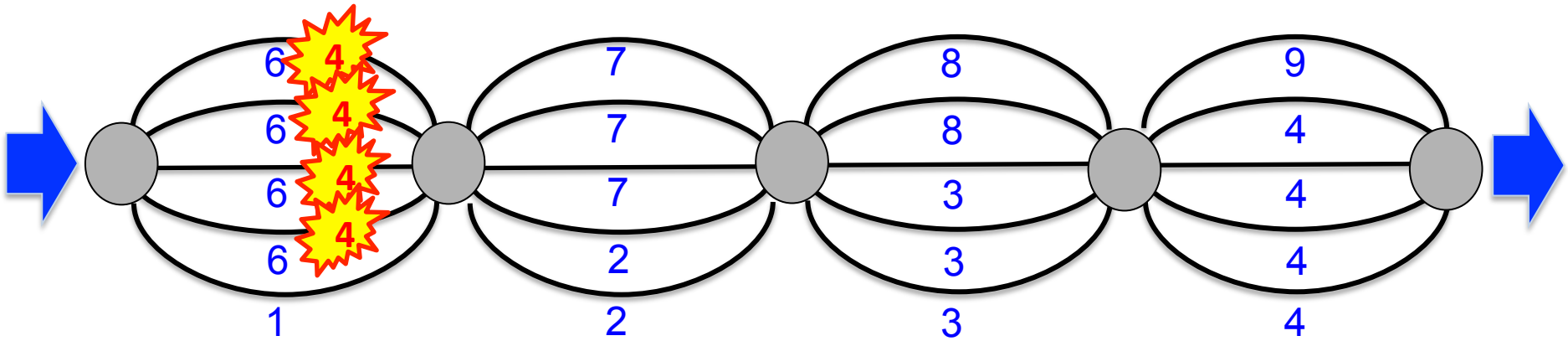
- With 0 attacks, capacity = 25
- With 1 attack, capacity = 16
- With 2 attacks, capacity = 9

Why simple priority lists don't work



- With 0 attacks, capacity = 25
- With 1 attack, capacity = 16
- With 2 attacks, capacity = 9
- With 3 attacks, capacity = 4

Why simple priority lists don't work



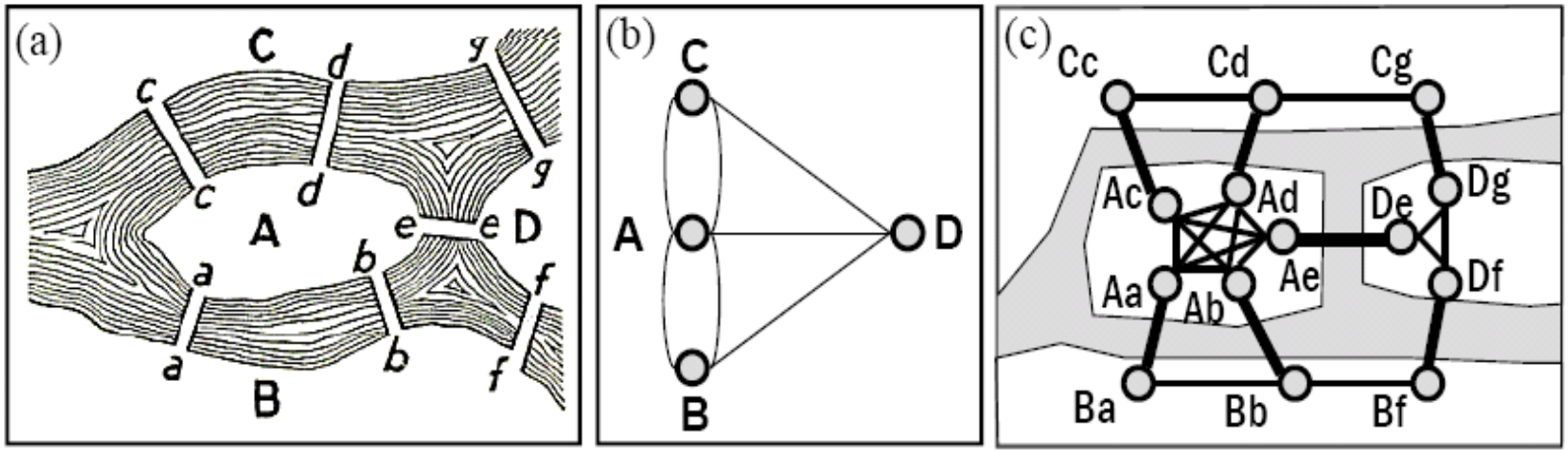
- With 0 attacks, capacity = 25
- With 1 attack, capacity = 16
- With 2 attacks, capacity = 9
- With 3 attacks, capacity = 4
- With 4 attacks, capacity = 1
- With 5 attacks, capacity = 0

~~Most vital arc?~~

Most vital sets of arcs!

Greedy heuristics (e.g., priority lists) won't work here.

The City of Königsberg



(a) from [Kraitchik 1942, pp. 209-211]

- This is not Euler's problem, however
- City administrators want to know which bridges are most important to protect against attack
- How much delay could commuters experience if terrorists destroyed the k-most-important bridges?

Reference: Alderson, D.L., Brown, G.G., Carlyle, W.M., and Wood, R.K., 2011, "Solving Defender-Attacker-Defender Models for Infrastructure Defense," in Operations Research, Computing, and Homeland Defense, R.K. Wood and R.F. Dell, editors, INFORMS, Hanover, MD, pp. 28-49.

Defender-Attacker-Defender Models

$$\min_{\mathbf{y} \in Y} f(\mathbf{y}) \qquad (\underline{\mathrm{D}})$$

$$\max_{\mathbf{x} \in X} \min_{\mathbf{y} \in Y(\mathbf{x})} f(\mathbf{y}) \qquad (\mathrm{A}\underline{\mathrm{D}})$$

$$\min_{\mathbf{w} \in W} \max_{\mathbf{x} \in X(\mathbf{w})} \min_{\mathbf{y} \in Y(\mathbf{x},\mathbf{w})} f(\mathbf{y}) \qquad (\mathrm{DA}\underline{\mathrm{D}})$$

Defender-Attacker-Defender Models

$$\min_{\mathbf{w}\in W} \max_{\mathbf{x}\in X(\mathbf{w})} \min_{\mathbf{y}\in Y(\mathbf{x},\mathbf{w})} f(\mathbf{y})$$

Defensive investments and budget constraints for the "defender"	Attacks limited by capability of the "attacker" and installed defenses	Optimal operation of the system after attack both defensive investment and attack

- 3-stage Stackelberg game: (1) Play, (2) observe and play, (3) observe and play. We seek w*.
- "Cost minimization" is broadly construed

DAD : Solving

$$\min_{\mathbf{w}\in W} \max_{\mathbf{x}\in X(\mathbf{w})} \min_{\mathbf{y}\in Y(\mathbf{x},\mathbf{w})} f(\mathbf{y})$$

- DAD is not easy to solve (e.g., exhaustive enumeration works in principle but is impractical)
- We know various decomposition methods work
- These methods looks like extensions of Benders decomposition

Modern Case Study: Bridges of SF Bay Area

- Industry validated model of nonlinear road congestion
- Asymmetric traffic demands (commuting patterns)
- Reconstitution times for each road or bridge segment
- 5-year time horizon
- Efficient computation

Ref: Alderson D, Brown G, Carlyle M, Wood K, "How to Assess the Value of Critical Infrastructure: A Worst-Case View of Risk and Its Implications for Defensive Investment,"2011 (in review)

Key Messages

1. Beware taking graph data at face value.
   - Sensitivity of results to errors in graph structure?
2. Consider system function (or purpose) when modeling system structure and behavior.
3. The "value" or "criticality" of an infrastructure component depends on its contribution to overall system function
   - deciding what is "critically important" to the system should be an output of analysis, not a required input
4. Game theoretic models of infrastructure resilience can be solved at practical scale for real problems
5. A "prioritized" list of targets is overly restrictive: it prohibits the efficient use of attack (or defense) resources.

Finding the Most Vital Arc in a Network: Lessons Learned from Infrastructure Defense

David Alderson, PhD
Operations Research Department
Naval Postgraduate School
dlalders@nps.edu

Graph Exploitation Workshop

MIT Lincoln Labs

August 2011