Threat Networks and Threatened Networks: Basic Principles and Practical Applications

Q1: What are the problems?

- Basic research in the science of network analysis to improve military and intelligence approaches to attacking and defending warfighting networks
- Development of improved tools for analysis of critical warfighting networks and for the disruption of opposing networks

Q2: Why care?

- Scientific: New Laws of Complex Networks
- Practical: example: Intentional attack vs. Random attack

Q3: What do we do("50 papers with 1700 citations" and 2 PDA softwares)?

Collaborators: L.A. Braunstein / S.V. Buldyrev / R. Cohen / S. Havlin
T. Kalisky /E. Lopez / G. Paul / S. Sreenivasan / H. E. Stanley
T. Tanizawa / Z. Wu, G. Li, Y. Chen, M. Kitsak

Thanks to: Drs. Goolsby, Walecki, and Shlesinger, Naval Sciences Board
GWOT Committee (National Academy of Sciences Representative)

Outline

1. *Background and Terminology*
2. *Network Immunization Strategies*
3. *Designing Networks Resilient Against Attack*
   - *A. Network Integrity (connected?)*
   - *B. Network Efficiency (cheap?)*
   - *C. Network Conductance (does it work?)*
4. *Network Design Tool： DEMO*
5. *Next Steps*

TWO TAKE HOME MSGS:

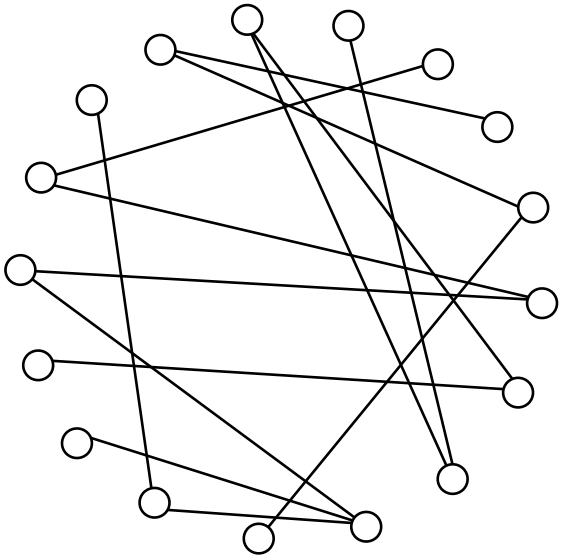
Statistical physics techniques are useful to

- Make progress in basic science of complex networks
- Determine practical network designs for real-world attack scenarios

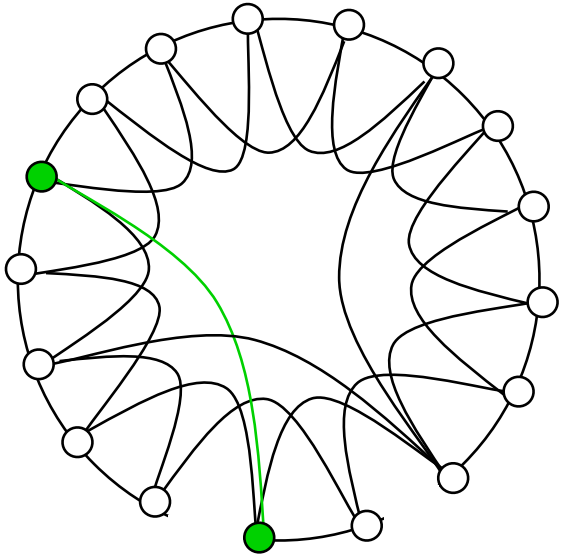
1. Terminology (Network Structure!)

3 families of networks:

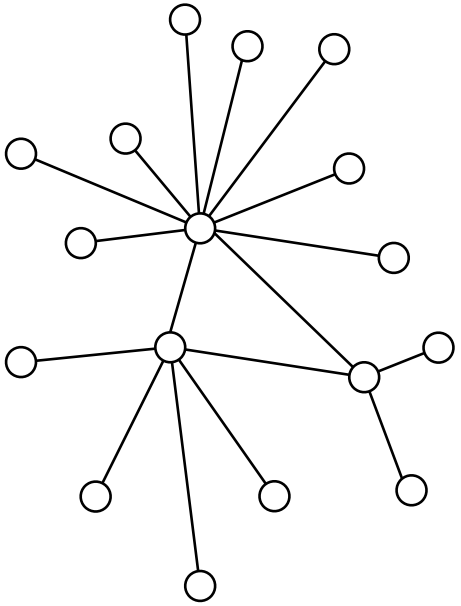
Erdös-Rényi
(Exponential tail)



Watts-Strogatz

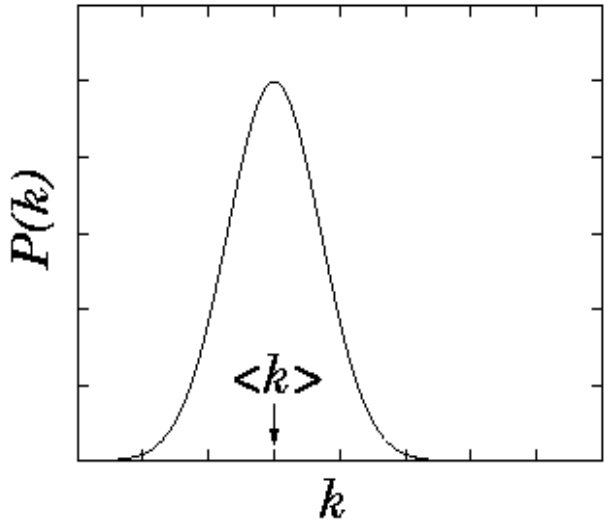


Scale-free
(Power law tail)

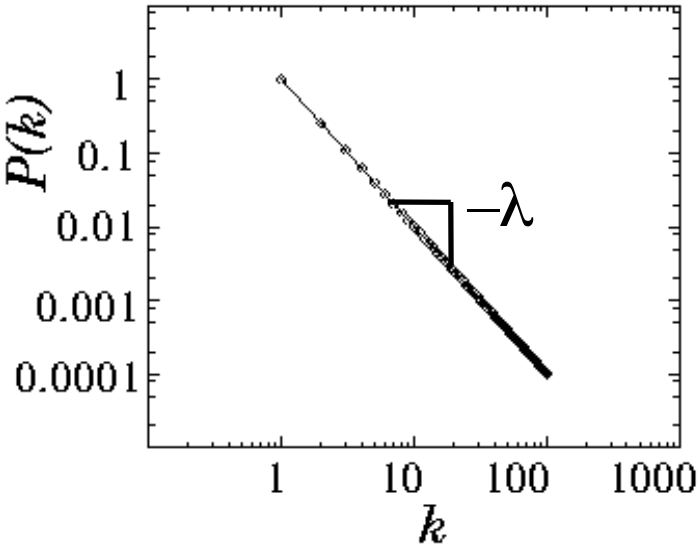


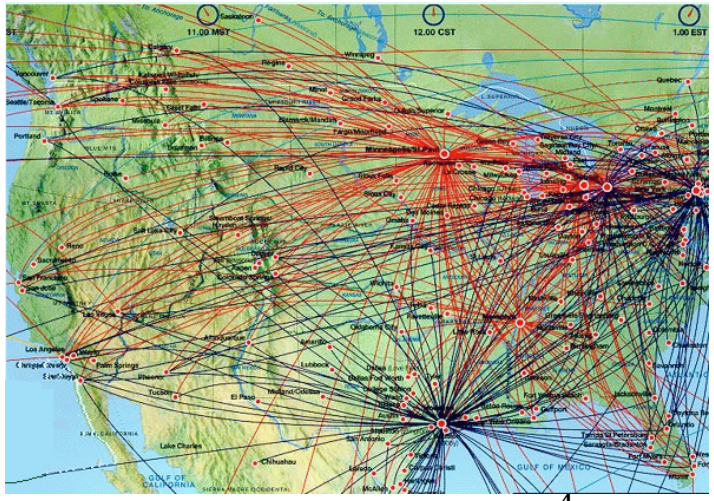
Example: Histogram of Number of nodes of degree *k*

Poisson distribution (exponential tail)



Scale-free distribution (power law tail)





Exponential Tail

Power Law Tail

New Result: How to Quantify Damage in social networks (with Borgatti)

Two ways to quantify network damage

- Percolation Theory:

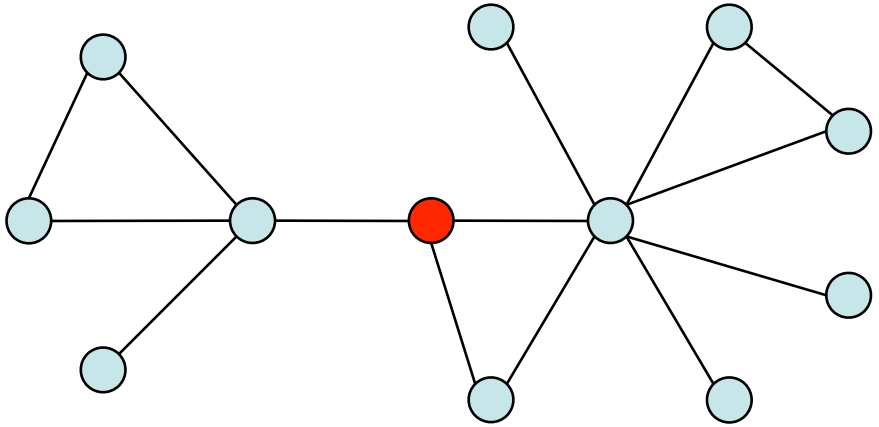
$$P_\infty = \frac{\text{size of largest set of connected nodes}}{\text{number of nodes in undamaged network}}$$

- Borgatti*:

$$C = \frac{\text{number of connected pairs of nodes}}{\text{number of pairs in undamaged network}}$$

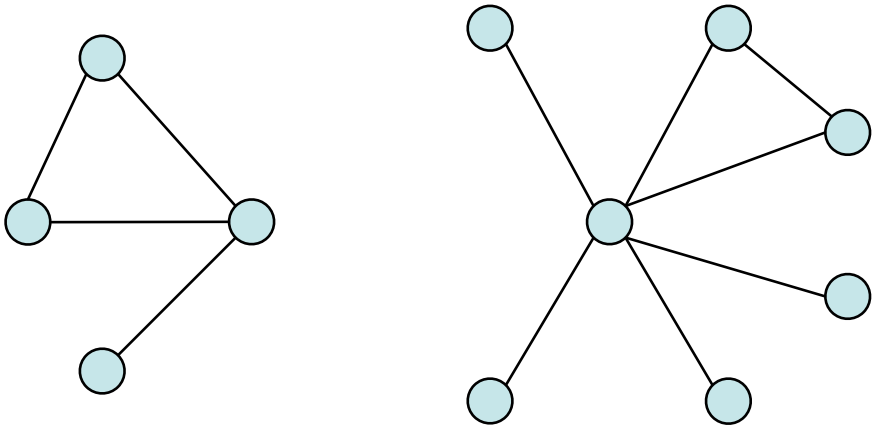
**Borgatti actually defines F=1-C*

Undamaged network



P_∞	C
1	1

Damaged network



P_∞	C
0.58	0.41

New Theoretical Result: Relation between Borgatti C and percolation

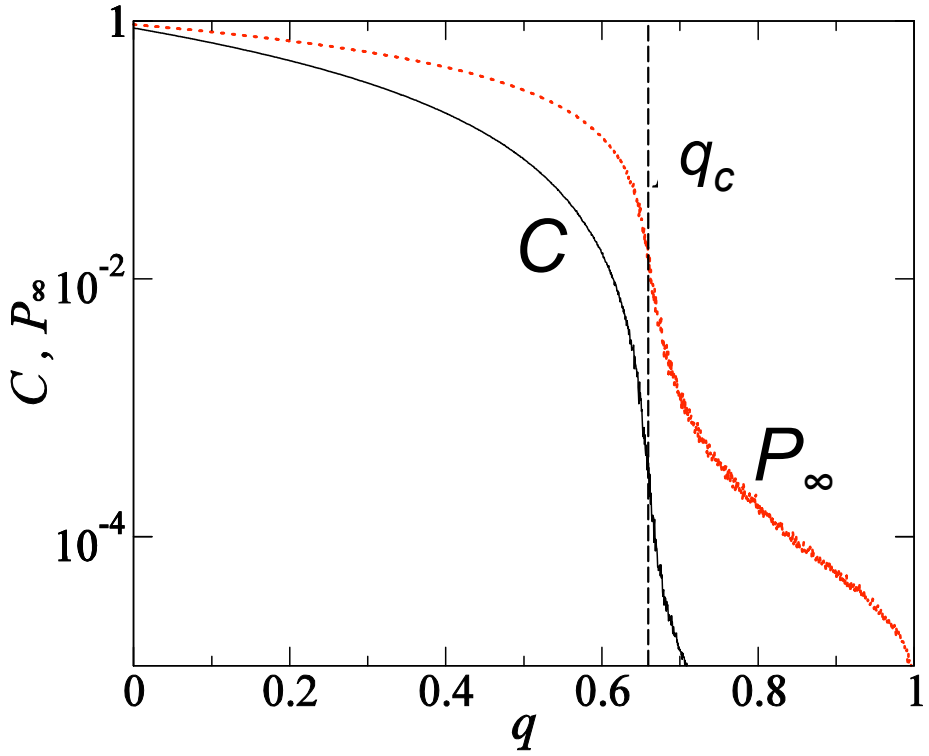
$$C = \frac{\sum_j N_j(N_j - 1)}{N(N-1)} \approx \frac{N_{\max}(N_{\max} - 1)}{N(N-1)} \approx \frac{N_{\max}^2}{N^2} = P_\infty^2$$

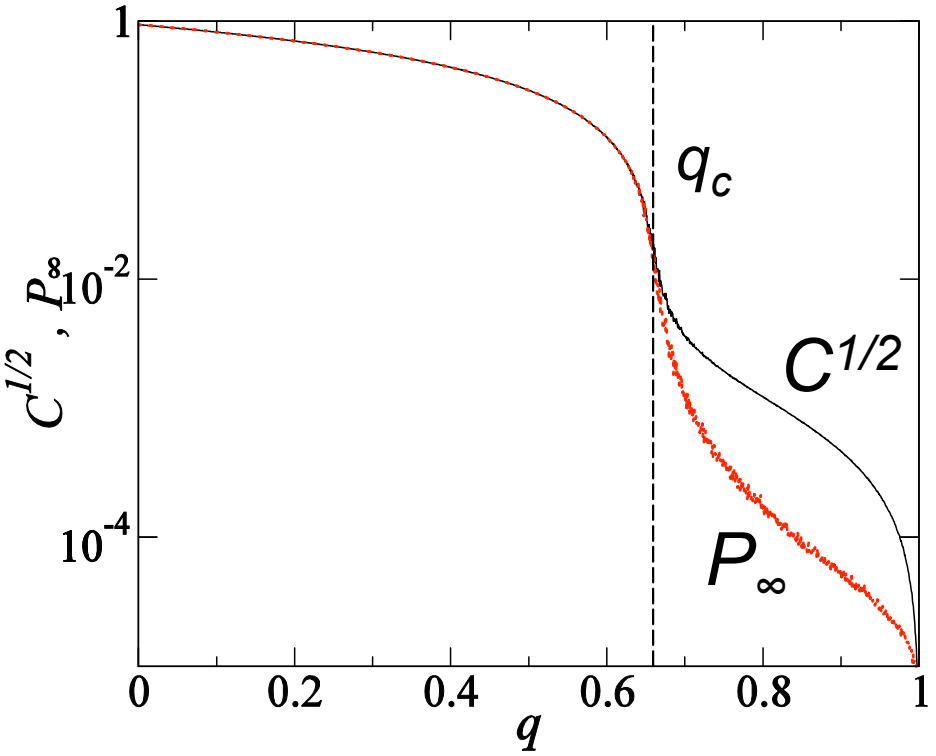
thus $P_\infty \approx C^{1/2}$

Next we test this result by simulations (N=200,000)

Test of New Theory Result: Simulations on Erdös-Rényi network

Strategy: random removal of nodes in a network of N=200,000 nodes



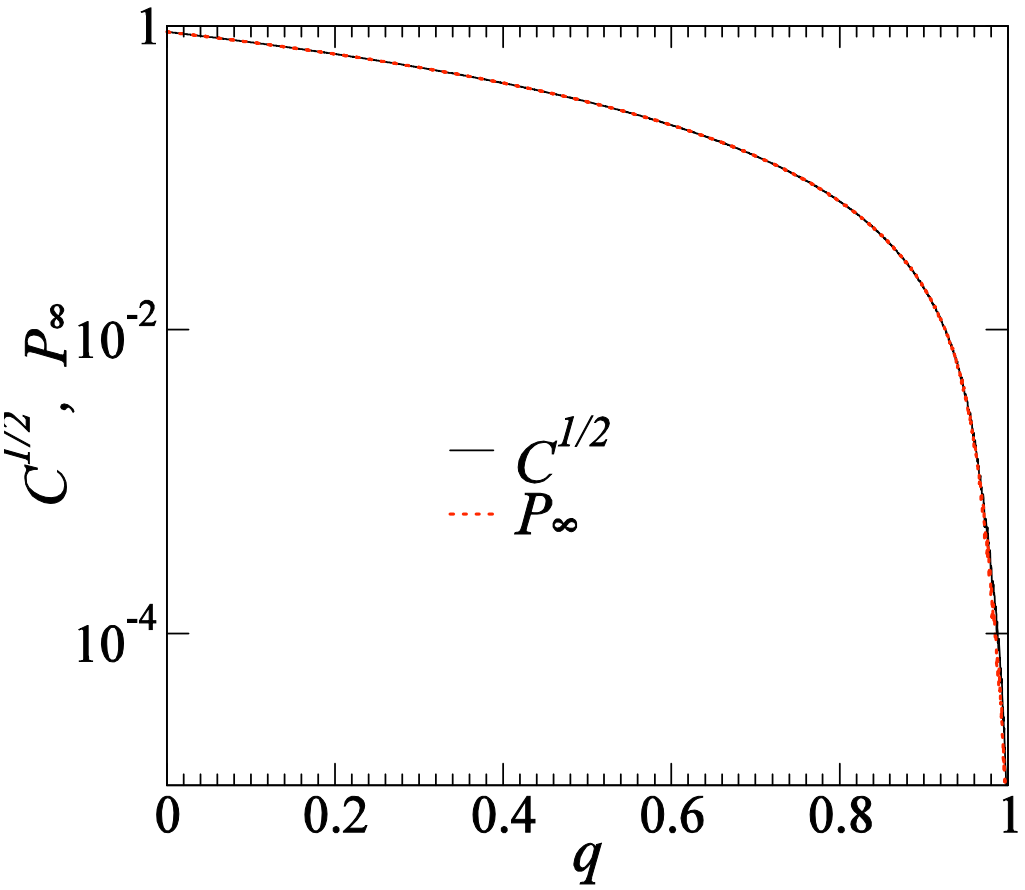


q = fraction of nodes removed from network

Test of New theory: a real social network (physics/sociol.collaboration)

Swedish network of workplaces

(N=300,000: from Sociologist Liljeros)



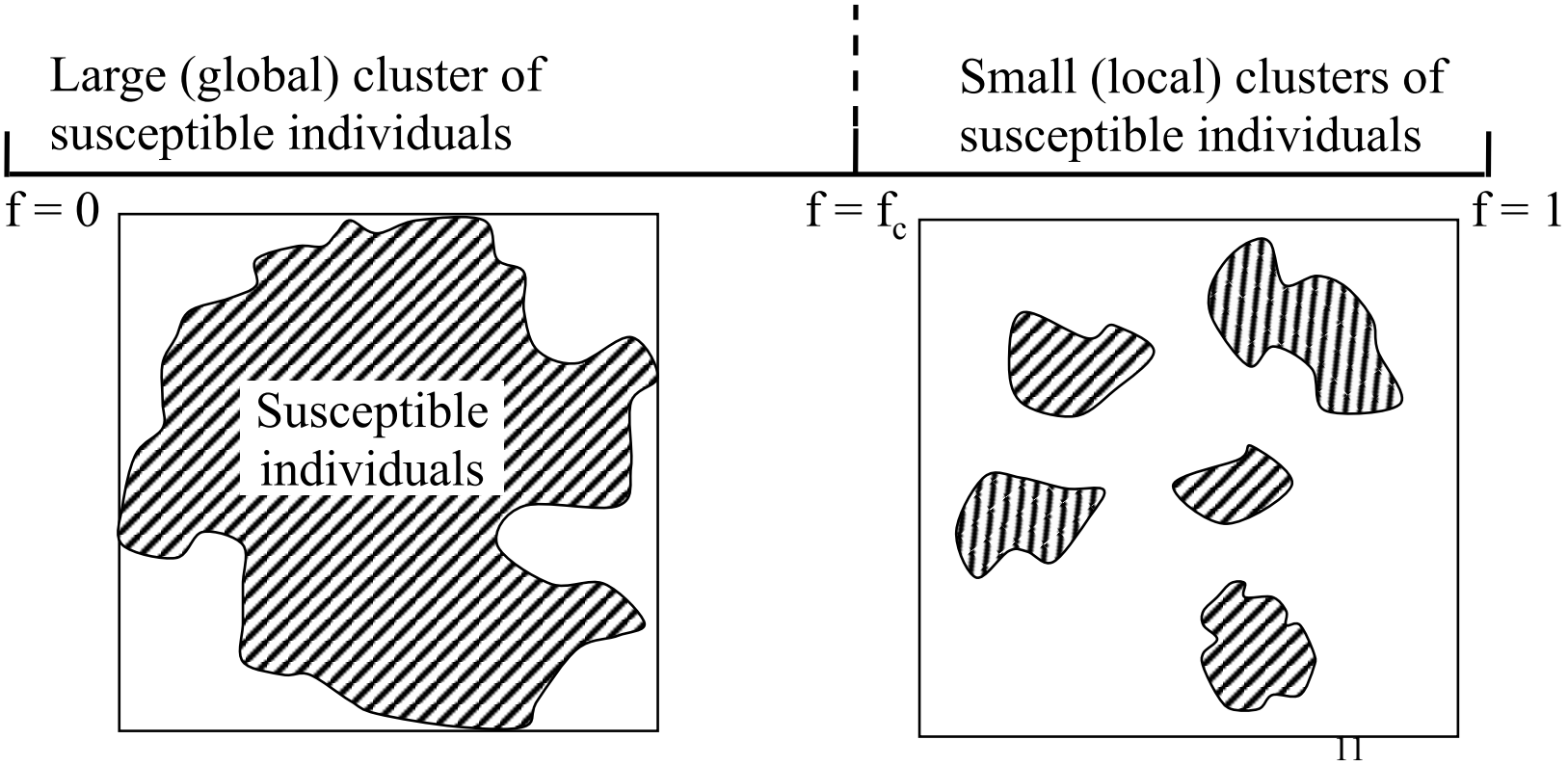
Impact of New Result

- Percolation theory can be used to infer results for *C, which is important in social network analysis.*

2. New Result: Network Immunization Strategies (analogy with attack: later)

Goal of efficient immunization strategy:

- Immunize at least a critical fraction f_c ("Immunization threshold") of the number of individuals so that only isolated clusters of susceptible individuals remain.
- Effective without detailed knowledge of the network.

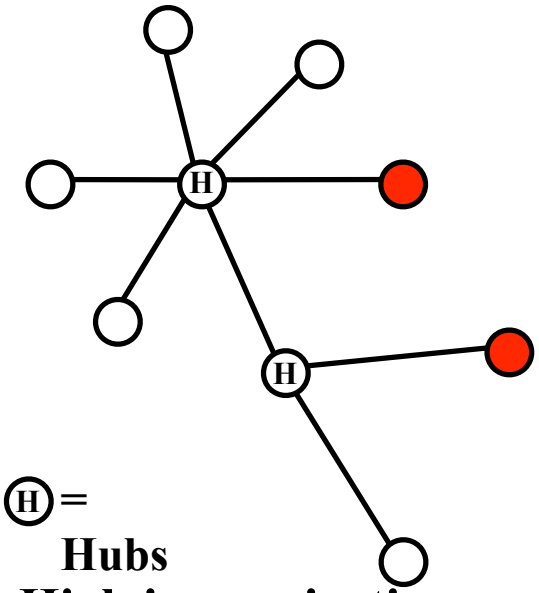


Three immunization strategies

Ex: Immunize 2 of the 9 nodes. What is immunization threshold?
i.e., what is chance to immunize ("KILL") the hubs?

Random:

2/9

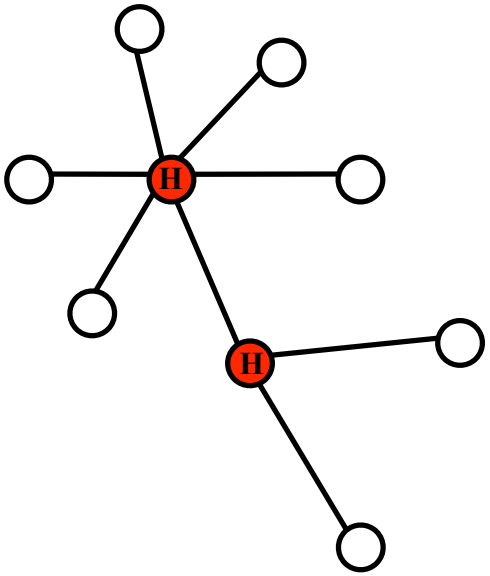


H = Hubs

- **High immunization threshold**
- **No prior network information needed**

Targeted:

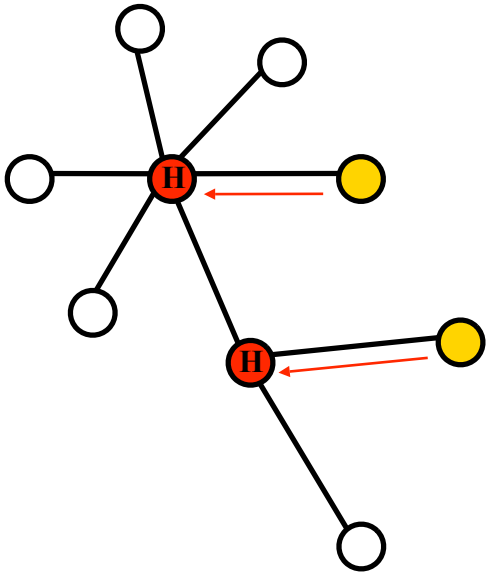
1



- **Low immunization threshold**
- **Knowledge of hubs (highly connected individuals) needed**

Acquaintance:

7/9

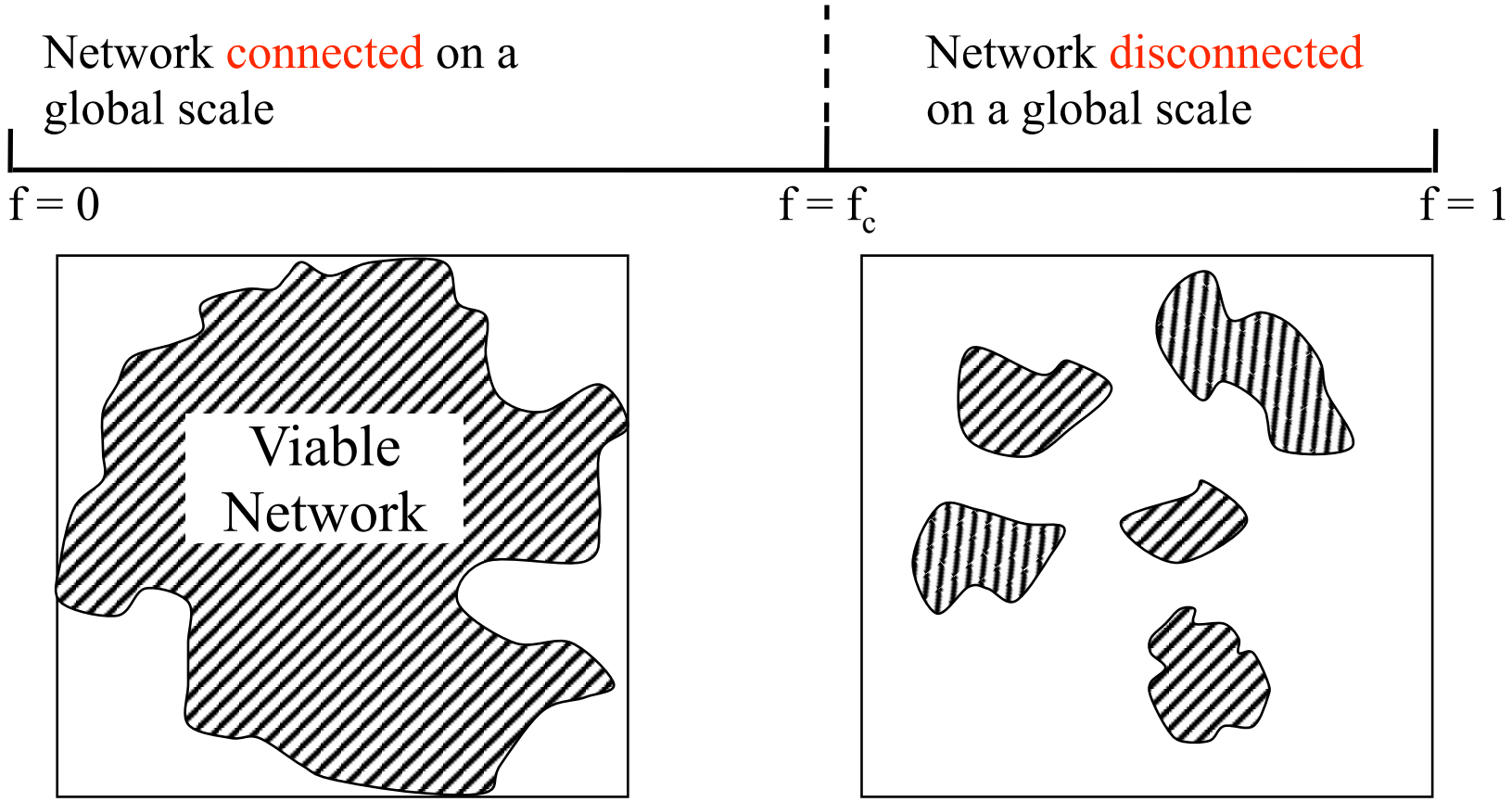


- **Low immunization threshold**
- **No prior network information needed**

3. Designing networks resilient against attack: New results on Network INTEGRITY

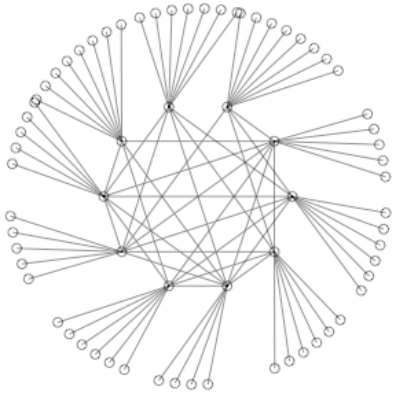
Immunization Goal: Attack the network (low immuniz. threshold f_c)

Resilience Goal: Protect the network (large breakup threshold f_c)



Question: how to design network to maintain integrity
Answer: not ER, not SF, but BINODAL

- Our maximally resilient network is more resilient than strictly scale-free or Erdös-Rényi networks to waves of attacks. (Surprisingly, our maximally resilient network has no scale-free attributes)
- Maximally resilient network has bimodal degree distribution
  - Fraction 1 – r of nodes have degree $k_1 = 1$
  - Fraction $r \approx 2\ p_t / p_r$ of nodes have higher degree: $k_2 = (\langle k \rangle - 1 + r) / r$



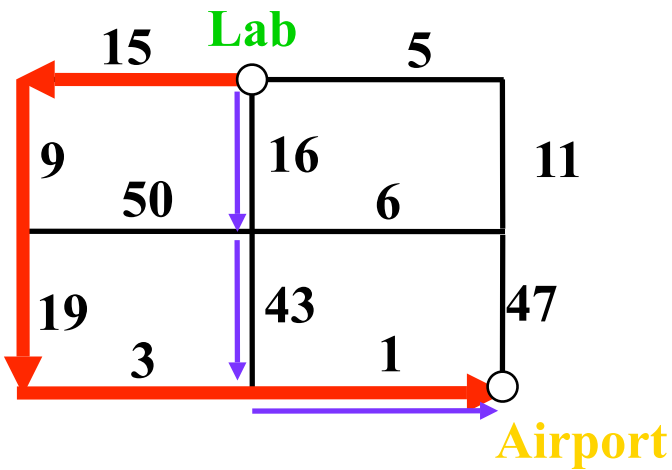
3B. New Results on Network EFFICIENCY

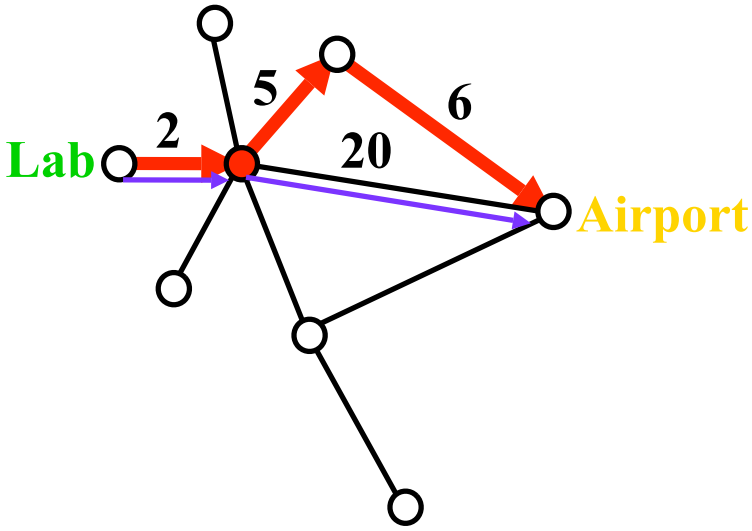
- Before: focus on protecting network integrity against attack (maintain CONNECTIVITY!)

- Now: focus on protecting network efficiency against attack (maintain OPTIMAL PATH!)

1. L.A. Braunstein, S.V. Buldyrev, R. Cohen, S. Havlin and H. E. Stanley Phys. Rev. Lett. (2003)
2. S. Sreenivasan, T. Kalisky, L.A. Braunstein, S.V. Buldyrev, S. Havlin and H. E. Stanley, Phys. Rev. E (2004)

Optimal Path: Minimize total "cost"





For this example:

Shortest path: 3 (cost = 60)
Optimal path: 5 (cost = 47)

Shortest path: 2 (cost = 22)
Optimal path: 3 (cost = 13)

Generally:

Shortest path = $N^{0.50}$
Optimal path = $N^{0.61}$
$N^{0.50} < N^{0.61}$
ex: $(10^6)^{0.50} < (10^6)^{0.61}$

Shortest path = Log N
Optimal path = $N^{1/3}$
$\log N << N^{1/3}$
ex: $N=10^6$, $\log 10^6 << (10^6)^{1/3}$

3C. New Results on Network Conductance

- Before: focus on protecting network INTEGRITY and EFFICIENCY.

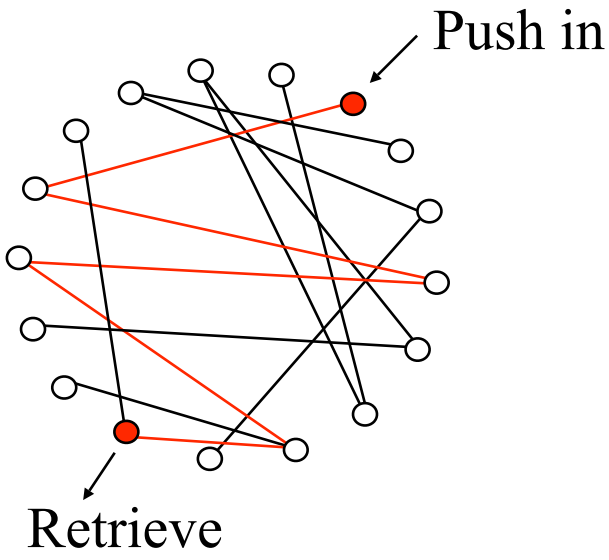
- Now: focus on maximizing network FLOW conductance (network structure which gives the highest flow conductance).

1. E. Lopez, S.V. Buldyrev, S. Havlin and H. E. Stanley (2004)
2. Z. Wu, E. Lopez, S. V. Buldyrev, L. A. Brauntein, S. Havlin and H. E. Stanley (2004)

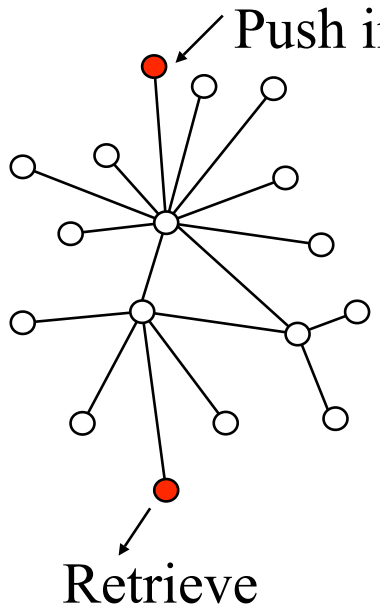
What is network conductance?

- "Push" goods, electrical current, cars, etc. into a network.
- Higher network conductance means higher flow of goods, etc., given equal "pushing".

"Old" (Erdös-Rényi)



"New" (scale-free)



How to maintain network flow?

- In a complex network, different links (or nodes) have different importance for transport.
- How to identify these important links ("superhighways") ?

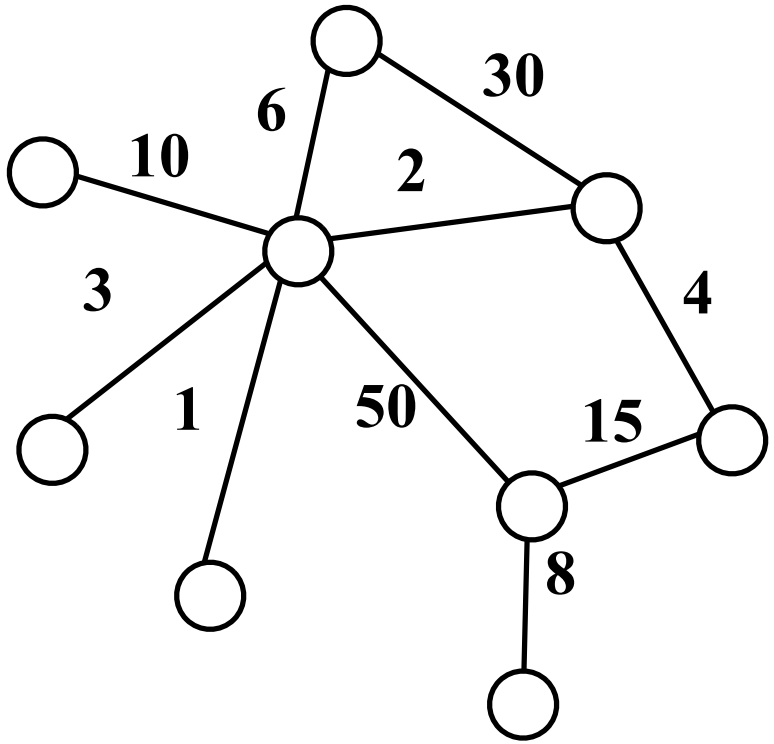
Motivation

•Identifying the superhighways and increasing their capacity enables us to improve transport.

•Also important for immunization. E.g., immunizing superhighways will reduce epidemics.

Z. Wu et al., Phys. Rev. Lett. **96**, 148702 (2006)

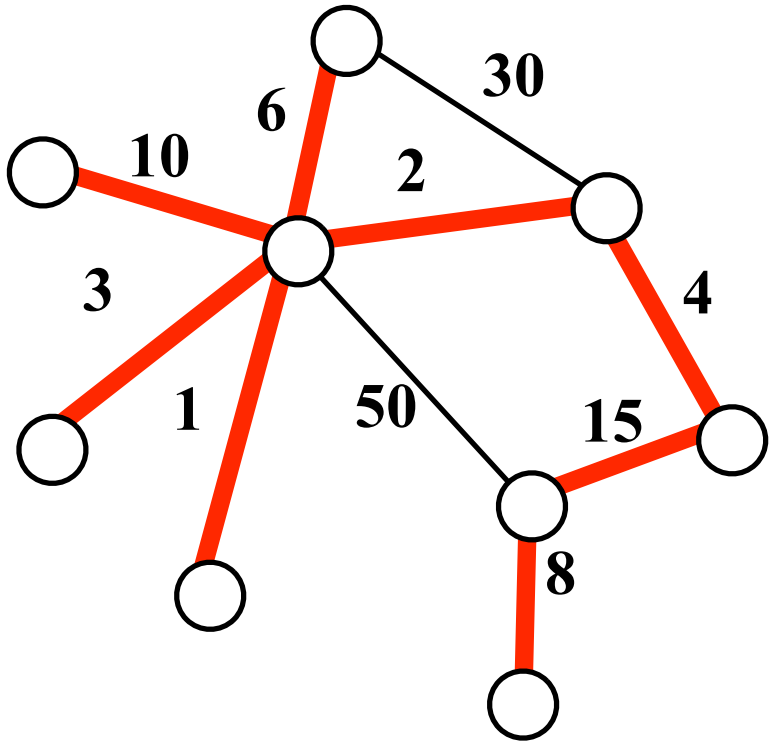
Def: Weighted networks



Barrat, Vespignani et al PNAS (2004)

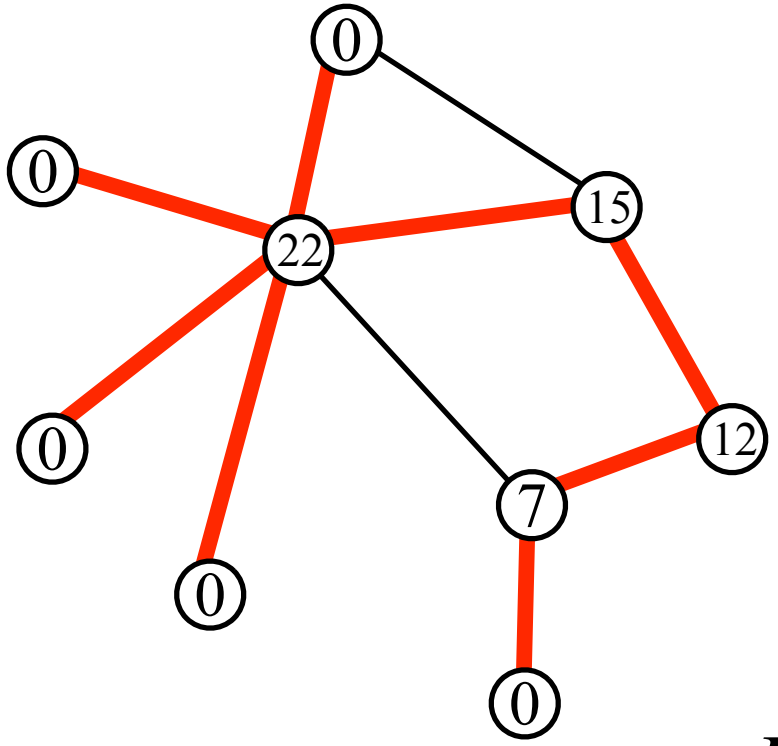
- Networks with weights, such as "cost", "time", "resistance" "bandwidth" etc. associated with links (or nodes).
- Many real networks such as world-wide airport network (WAN), *E Coli*. metabolic network etc. are weighted networks.
- Many dynamic processes are carried on weighted networks.

Def: Minimum Spanning Tree (MST)



- Tree = loopless graph
- MST = Tree with the lowest total weight which connects all nodes (red).
- MST used for network design.
- Fact: Most flow paths between pairs of nodes lie on the MST.

Definition: Betweenness centrality BC



- BC = Number of times a node (or link) is used by the set of all shortest paths between all pairs of nodes - betweenness centrality.*
- Measure the frequency of a node being used by flow. NOT unimodal. Rather:

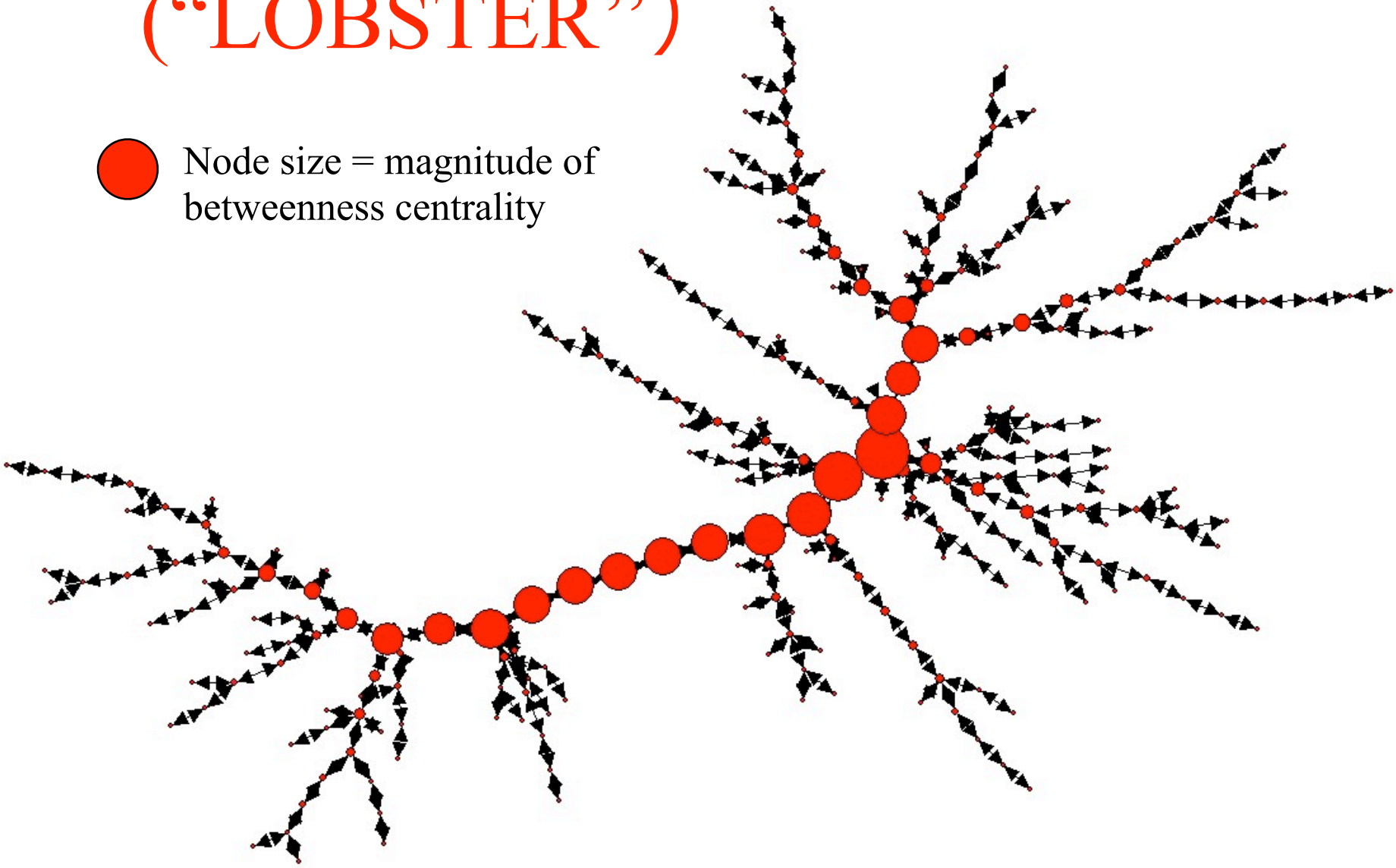
$$P_{\mathrm{MST}}(BC) \sim (BC)^{-\delta} \qquad \delta \approx 2^{**}$$

* Freeman, L. C. Sociometry, 40:35-41 (1977)

** D.-H. Kim et al., Phys. Rev. E **70**, 046126 (2004).
K.-I. Goh et al., Phys. Rev. E **72**, 017102 (2005).

Ex. MST of weighted ER network ("LOBSTER")

Node size = magnitude of betweenness centrality



Def: Incipient percolation cluster (IIC)

- Imagine we remove link with highest weight first, the second highest weight next, etc… Initially the largest connected component scales as N. However eventually the cluster fragments into small pieces and this statement is no longer true. We are at the "percolation threshold".

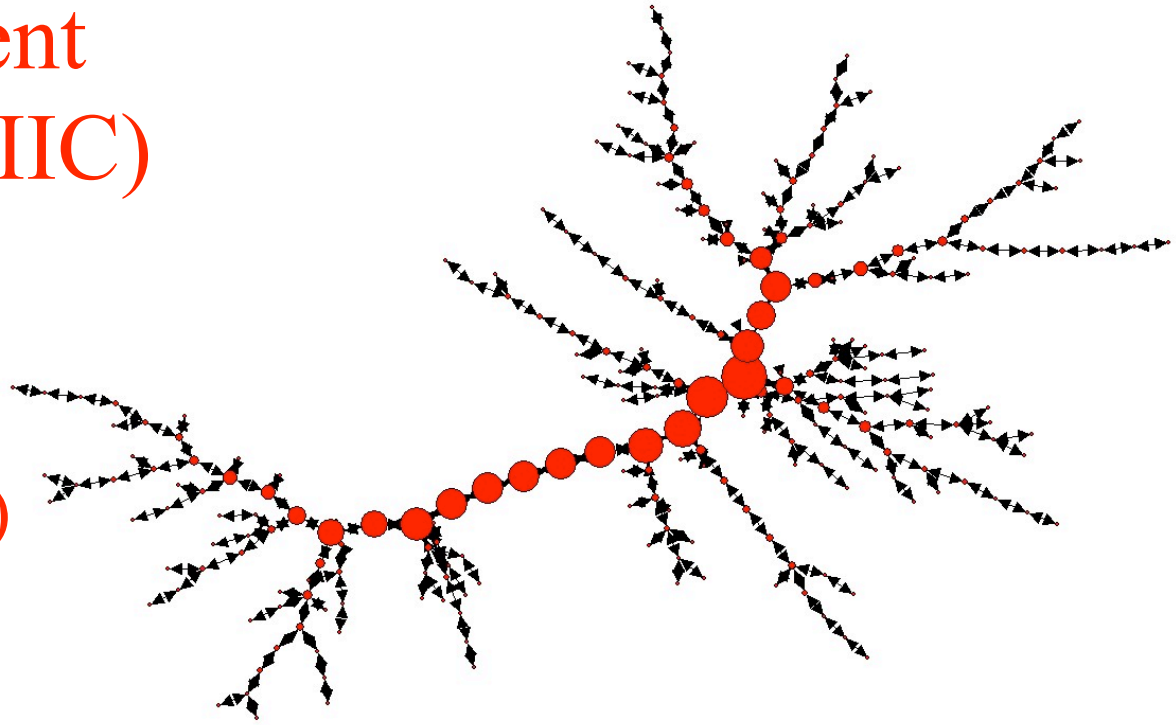
- IIC = largest component at percolation threshold, which occurs when the normalized size of largest connected component S/N approaches zero

- The IIC is a subset of the MST

Example: Incipient Infinite Cluster (IIC)

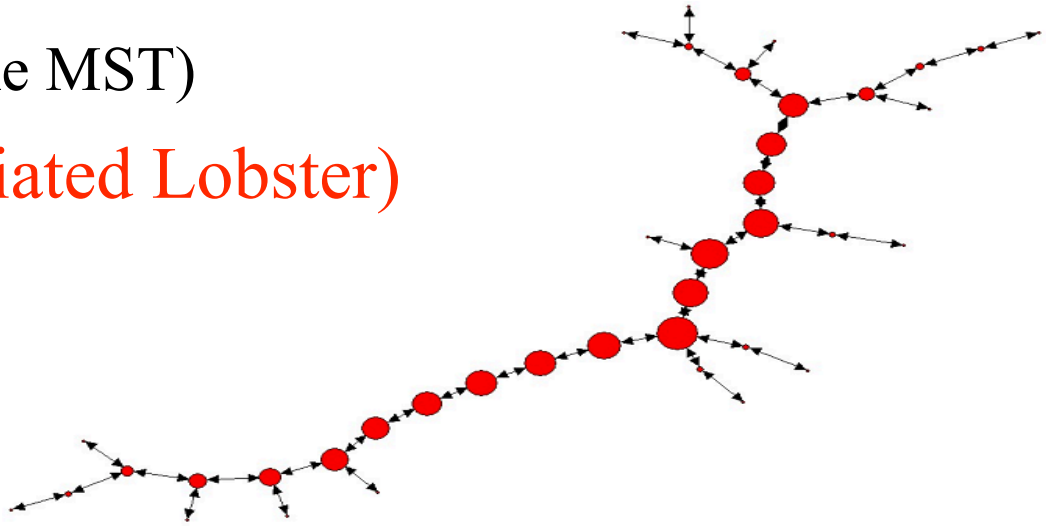
MST:

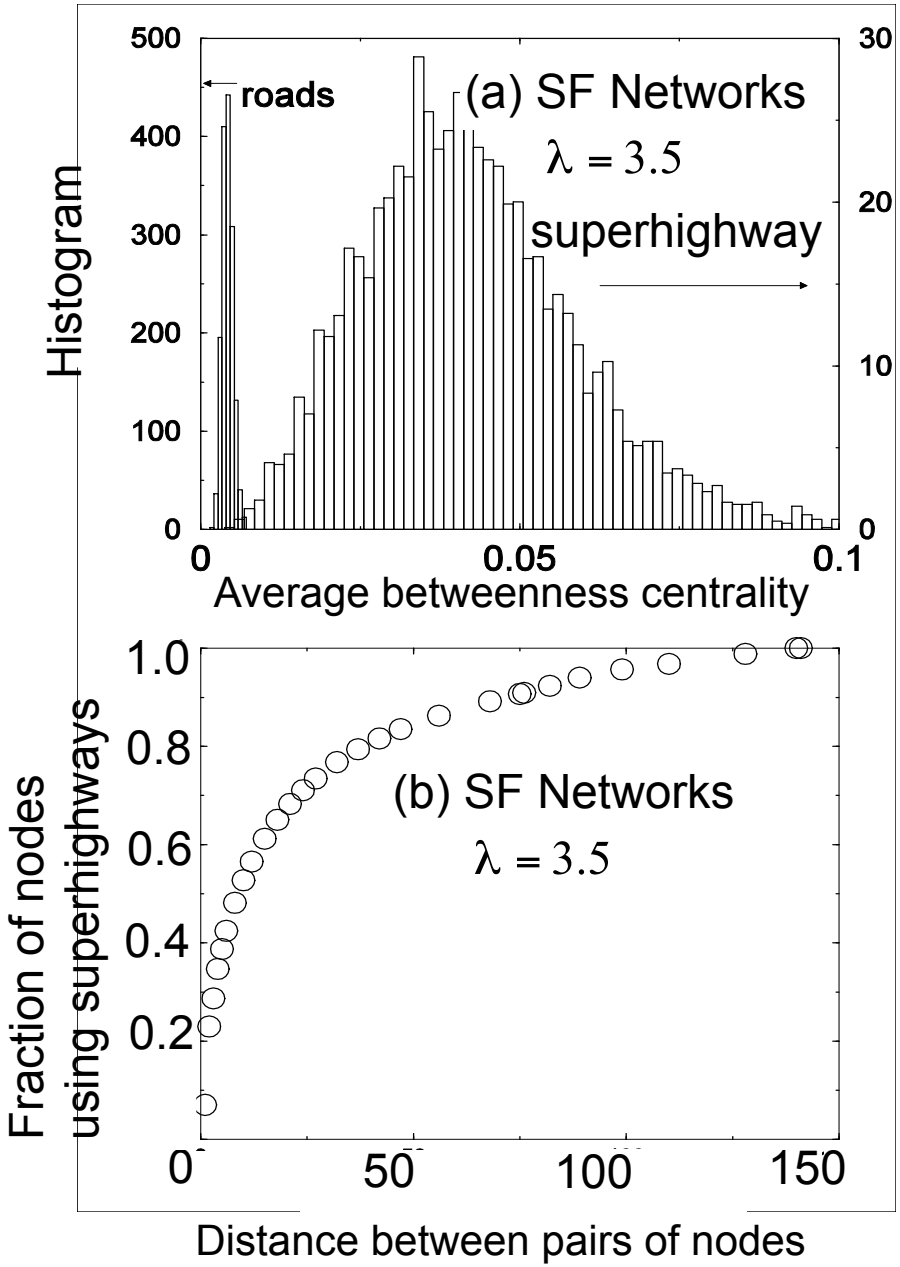
Superhighways and Roads (lobster)



I I C: (IIC = a subset of the MST)

Superhighways (Emaciated Lobster)





Flow on roads vs. superhighways (N=8192):

a) IIC has nodes of larger betweenness centrality!!!

(Def: Ave = BC/N^2)

b) Nodes farther apart more likely to use superhighways than roads!!!

Application of new results: improve network flow

Comparison between two strategies:

A: increase capacity of all links on superhighways.

B: increase capacity of the links with highest betweenness centrality in MST (same number of links as A).

First, we study flow in weighted random network.

Result: Strategy A yields higher increase in total flow.

Next, we study maximum flow problem, where each link of the network has an upper bound capacity (e.g. traffic, computer science).

Result: Strategy A yields higher increase in total flow.

Summary of 4 new results on flow:

Wu, Braunstein, Havlin, Stanley, PRL, 96, 148702 (2006)

- MST can be partitioned into superhighways which carry most of the traffic and roads with less traffic.
- We find the superhighways to be the IIC (largest connected component at the percolation threshold). The number of links in superhighways is of order $N^{2/3}$ -- a zero fraction of the network!!
- Increasing the capacity of the superhighways improves transport more than increasing the capacity of the links with high BC (SURPRISE).
- Immunizing the superhighways reduces epidemics more than immunizing the links with high BC (SURPRISE).

5. Next Steps

- Improve Network Design Tool （demo
- Consider mobile nodes in network
- Add network conductance as quantity to be optimized
- Add other real world network attributes/ constraints
- 10^6 node networks (needed for reliable predictions)

Summary so far:

Statistical physics techniques can be used to

- Make progress in basic science of networks.
- Determine optimal network designs against real-world scenarios.

4. NEW RESULT：Build "Network Design Tool" （demo）

Motivation:

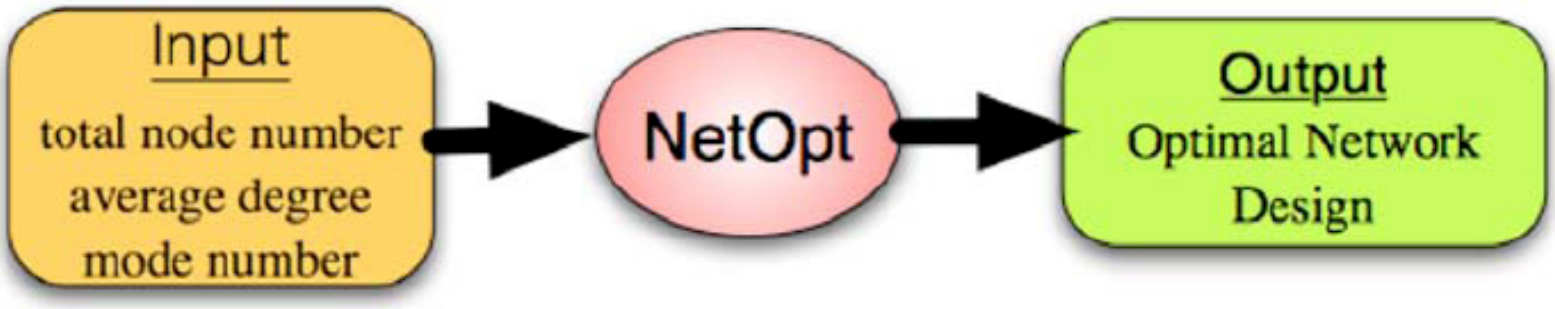
- Goal: make our work accessible for broader set of users.
- Input: attack parameters, network constraints,…
- Output: optimal design

NetOpt and NetAttack （PDA）

I. The Purpose of NetOpt and NetAttack

NetOpt:

NetOpt helps you to determine the network structure as robust as possible against both random failure of nodes and targeted attacks from the outside when the fundamental parameters such as the total number of nodes, the average number of links of each node, and the number of "modes" in the degree distribution are given.

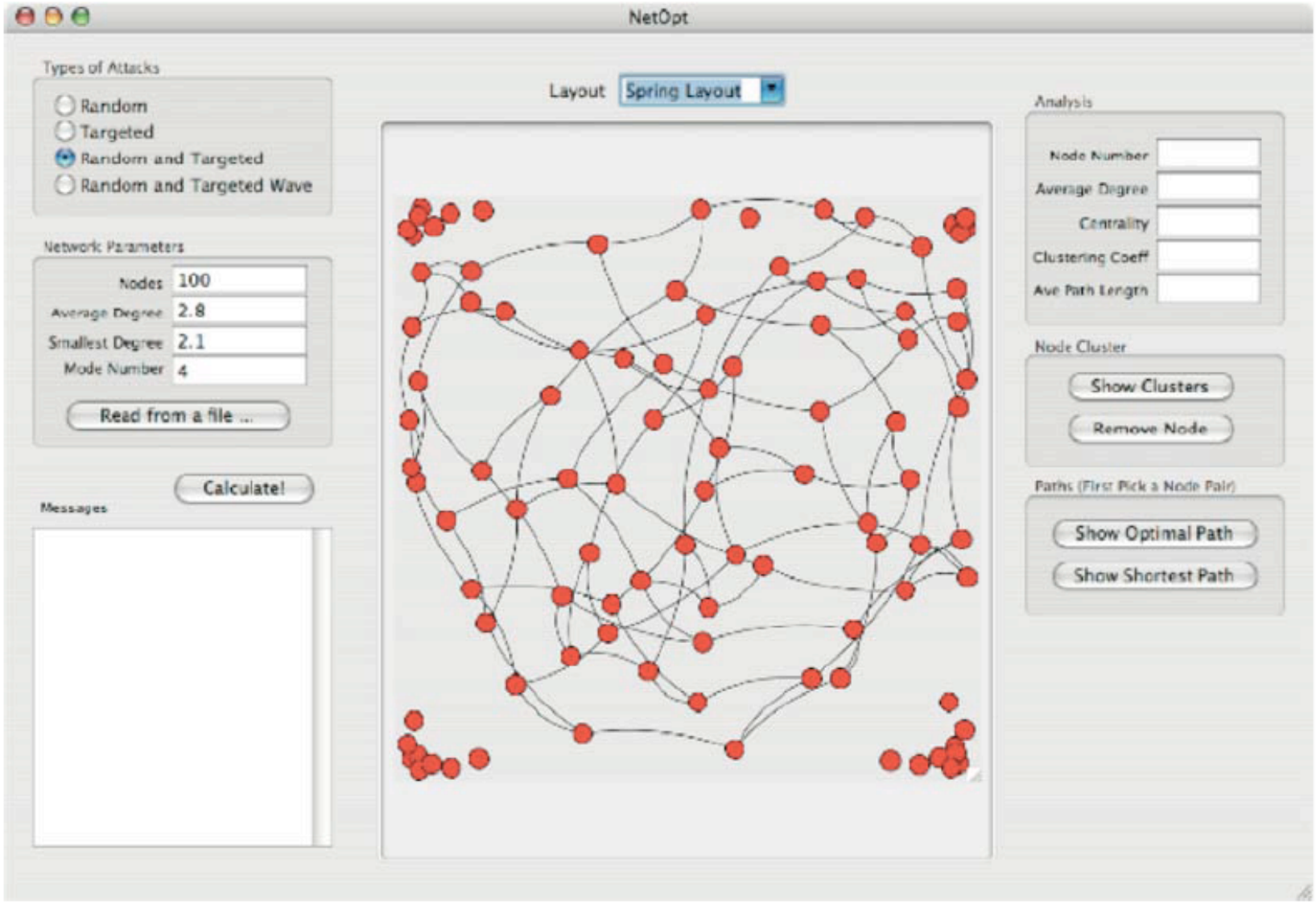


NetAttack:

NetAttack helps you to choose the best strategy for the disintegration of a given network even if the global structure of the network is not known. From the local information of the network structure, NetAttack gives the possible network structure and chooses the best strategy for the attack.

II. Features:

Simple input parameters with Graphic User Interface:



- *Real-time visualization of the network structure*
  - When you change the network parameters, you can see the new structure of the network almost instantaneously.
- *Various types of network layout*
  - You can choose the most suitable layout for network visualization from various types of network layout, such as "Spring layout", "Tree layout", "Circular layout" and so on.
- *Portability*
  - NetOpt/NetAttack is written in JAVA and you can run it equally on various types of platform such as Windows, Mac OS, and UNIX/Linux.
- *Including the latest results of network research*
  - Since NetOpt/NetAttack is made by one of the most active network research groups in the world, NetOpt/NetAttack is always updated including the latest results in this field.

III. What do we have now?

- ✓Theory for network robustness against various types of attacks (developed at Boston Univ)
  - ➢ *We also have other important results such as the theory for the crossover behavior of the optimal path in strongly disordered networks.*

- ✓Prototype of the application
  - ➢ *See the GUI example above.*

IV. What are we going to do?

- Include our latest studies for the robustness of multi-modal networks
- Include our latest studies for the behavior of optimal paths in weakly and strongly disordered networks
- Refine the appearance of the network visualization

$$F = 1 - \frac{\sum_{j=1}^{i} N_j (N_j - 1)}{N(N-1)} \equiv 1 - C$$

N_j – number of nodes in cluster j

i - number of clusters

C - connectivity measure

C=1 all nodes are connected ; fragmentaion F=0

C=0 all nodes are disconnected: F=1

In physics we use P_∞ the fraction of the largest cluster

Theory for Centrality Distribution

For IIC inside the MST:

$n_\ell : \ell^3$ for network at criticality

n_ℓ is number of nodes in MST within ℓ

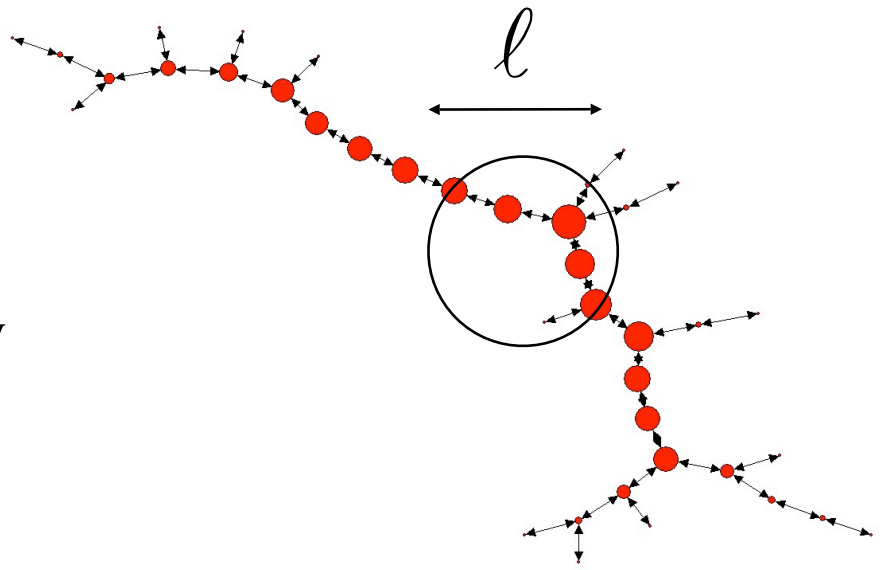
$s_\ell \sim \ell^2$ for nodes in the IIC

Thus the number of nodes with centrality larger than n_ℓ is

$$m(C > n_\ell) : \; n_\ell^{1/3} \frac{S}{s_\ell} \approx \frac{n_\ell^{1/3}}{n_\ell^{2/3}} S \sim n_\ell^{-1/3}$$

for all ℓ due to self-similarity. Thus,

$$p_{IIC}(C) : \; C^{-4/3}$$



For the MST:

$$m(C > n_\ell) \sim n_\ell^{1/3} \frac{N}{n_\ell} \approx \frac{n_\ell^{1/3}}{n_\ell} N \sim n_\ell^{-2/3}$$ Thus,

$$p_{MST}(C) = \frac{dm}{dn_\ell} \sim n_\ell^{-5/3} \sim C^{-5/3}$$

Good agreement with simulations!

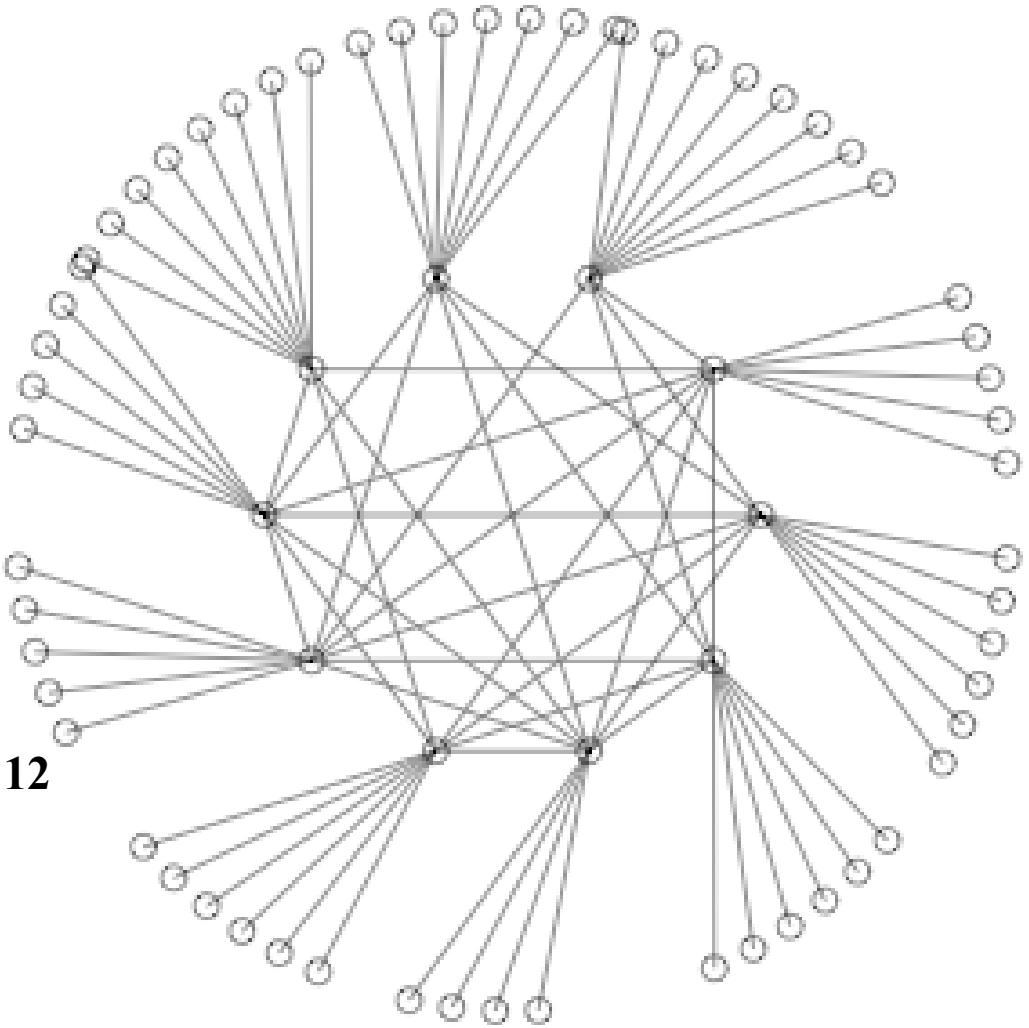
Maximally Resilient Network

Specific example:

Given: N = 100,
‹k› = 2.1,
$p_t / p_r = 0.05$

Optimal design is:
r = 2 * 0.05 = 0.1
10 "hubs" of degree

$k_2 = (2.1 - 1 + 0.1) / 0.1 = 12$

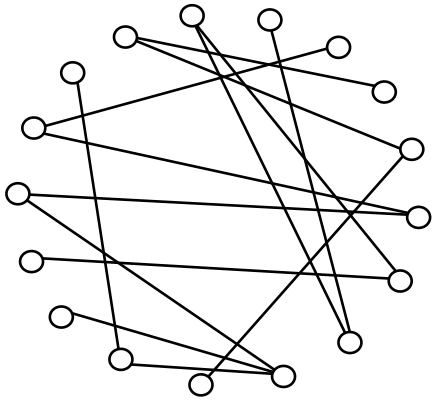


3A. New Results on Network Integrity: Realistic model

Multiple waves of alternating

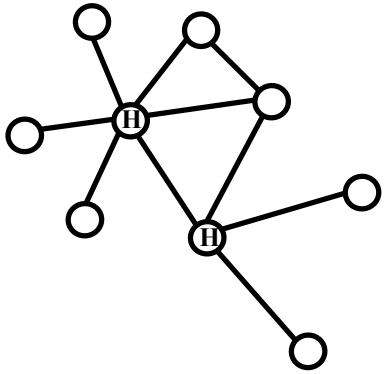
- Targeted attack with probability p_t
- Random attack/failures with probability p_r

Erdös-Rényi



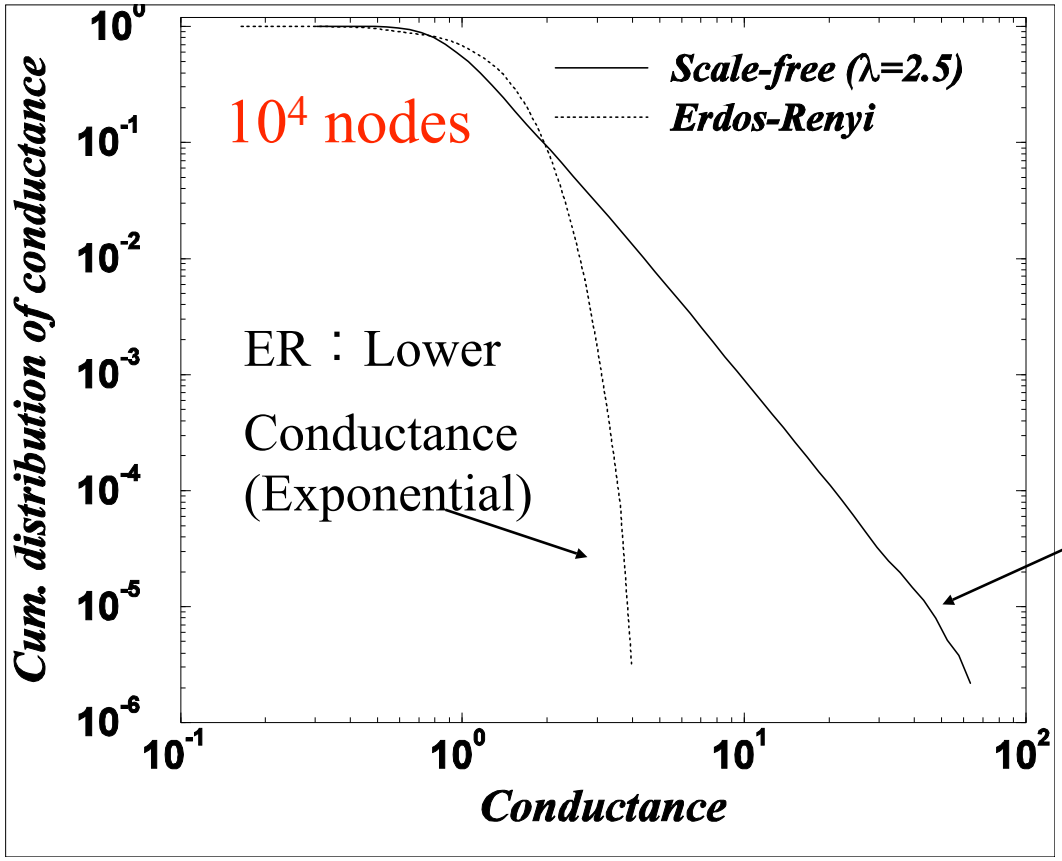
- Random attack:
  must remove ≈ 50% to destroy
- Targeted attack:
  must remove ≈ 50% to destroy

scale-free



- Random attack:
  must remove ≈ 99% to destroy
- Targeted attack:
  must remove ≈ 1% to destroy

RESULT: Scale-free networks have MUCH higher flow conductance than Erdos-Renyi networks!!!



Real world example of scale-free network: Internet Network

Faloutsos et. al., 1999