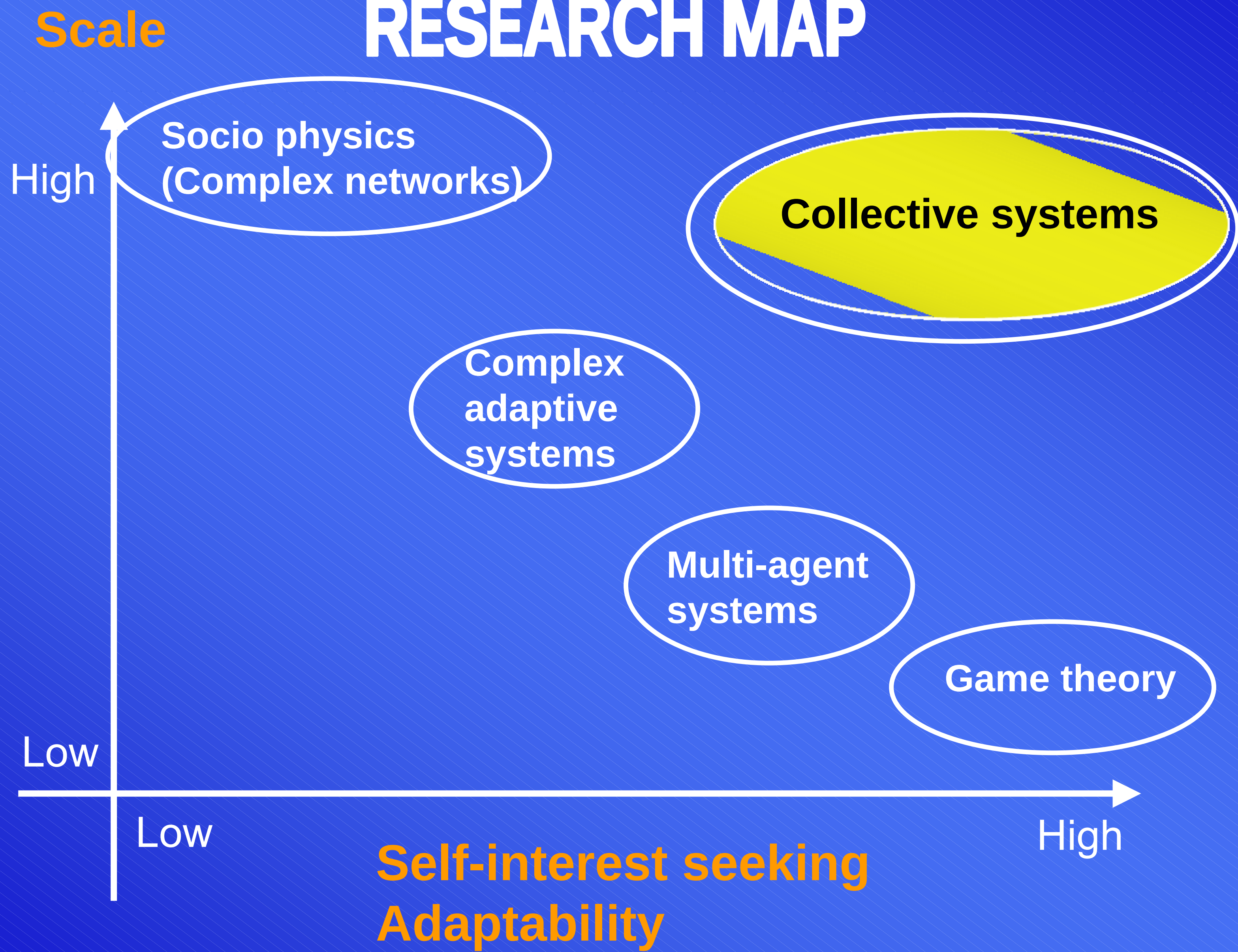
Optimization in Networked Systems

Akira Namatame(生天目　章)
Seung-Youp Shin(申　昇燁)
National Defense Academy of Japan
www.nda.ac.jp/~nama

RESEARCH MAP

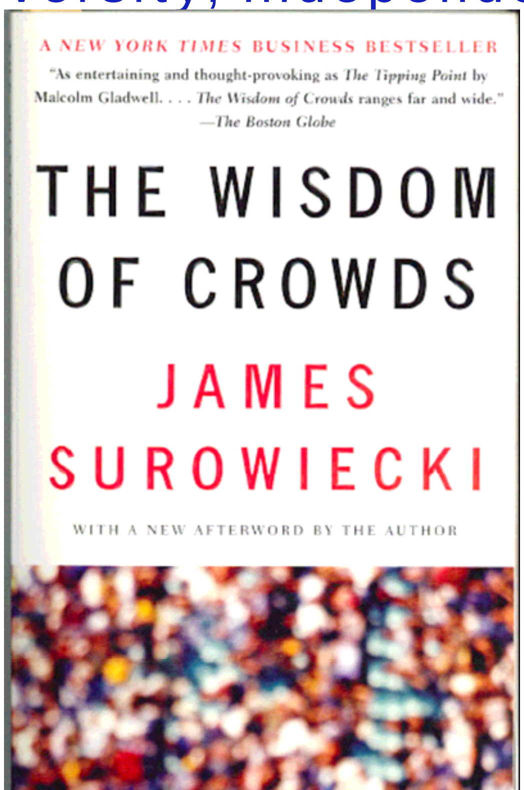


Wisdom of Crowds: Collective Intelligence

A large collection of people are smarter than an elite few.

In “the Wisdom of Crowds”, Surowiecki,(2004) suggests new insights regarding how our social and economic activities should be organized.

: The **wisdom of crowds** emerges only under the right conditions (diversity, independence, etc)



Phase Transition in Collective Behavior

- Crowds are wise, but are *also often foolish.*

Then under what mechanism can we improve the performance of collective systems?

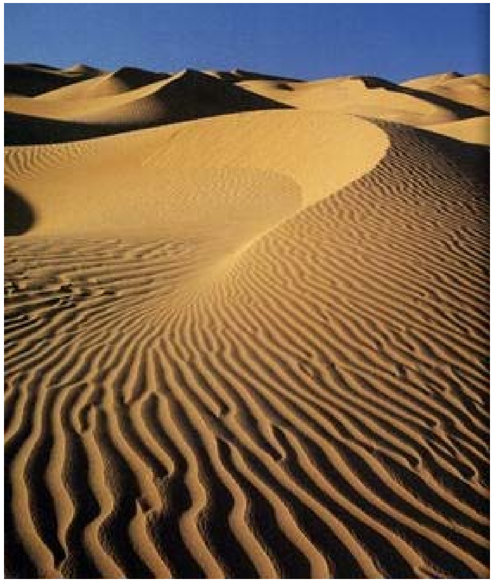
:The way of interaction, the network topology, plays a crucial role

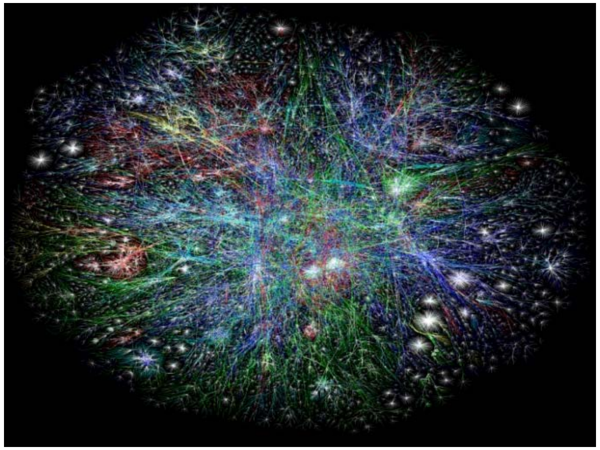
Emergence by Nature

- Emergence by nature (empirical view)
- View emergence as an “innate property” of natural systems

“*Systems self-organize into a complex state, poised between predictable cyclic behavior and unpredictable chaos*”

: Inspires research to discover and explain emergent behaviors





Internet

Emergence by Design

- Emergence by design strategies (operational view)

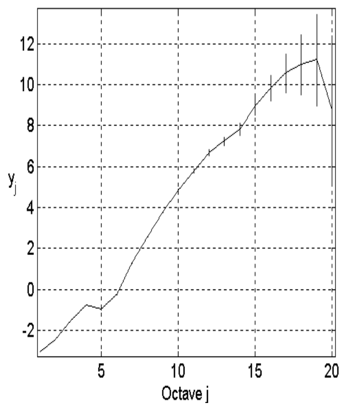
"System-wide behavior emerges from interactions among individual elements"

: Some researchers view emergence as a property that is "designed" into systems

: Inspire research into design techniques to induce desired emergent behaviors

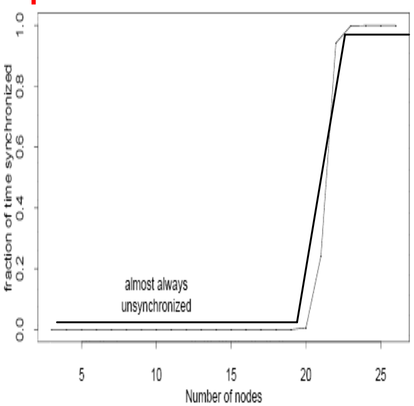
Emergence by Design: Illusion of Control?

self-similarities



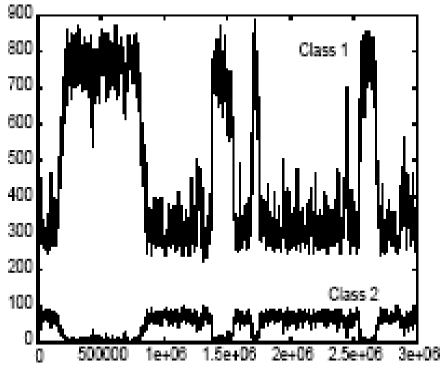
Internet throughput

phase transitions

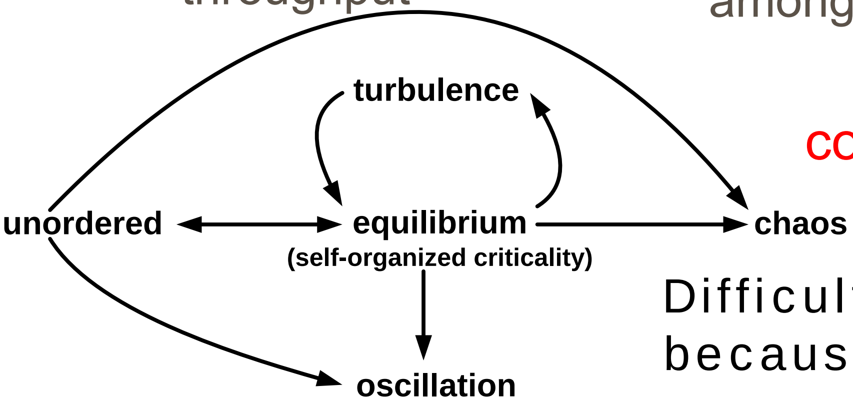


synchronization among routers

Meta-stabilities



distribution of call types in wireless cells



congestion collapse

Difficult to predict and control because of phase-transitional behavior

high load ➡ and the network topology play crucial roles

Emergence by Design: Evolutionary Optimization (1)

- Design of Communication Networks

  Tradeoff between congestion and network design cost

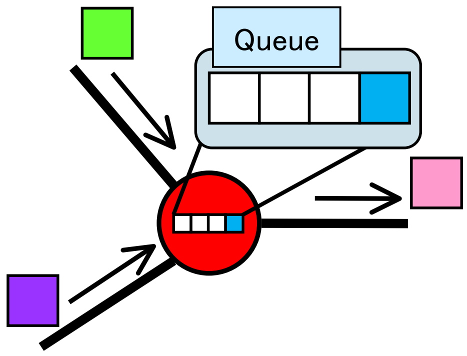
- Diffusion of Innovation

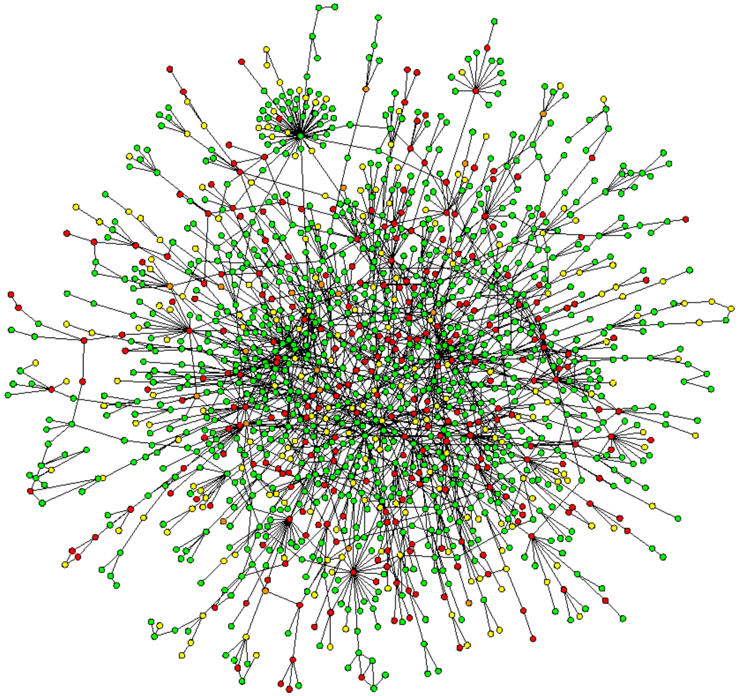
- Consensus (synchronization) in Networked Systems

A Network Flow Model

- Packet generation
  - Packet is generated at random with some rates
- Each nodes process one packet per time
- Each node has a queue to store undelivered packets
  - Routing: Shortest path
- Traffic congestion is determined by node betweenness

  : total shortest paths through the node

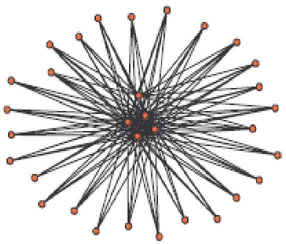




Optimized Networks: Minimizing congestion

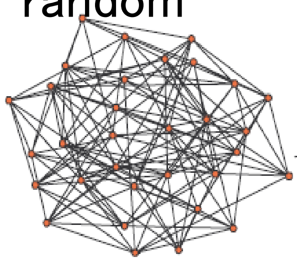
- Network size: 32 nodes
- Fixed number of links

- Optimal network
  - Star network: packet generation rate is small
  - Random network: packet generation rate is high

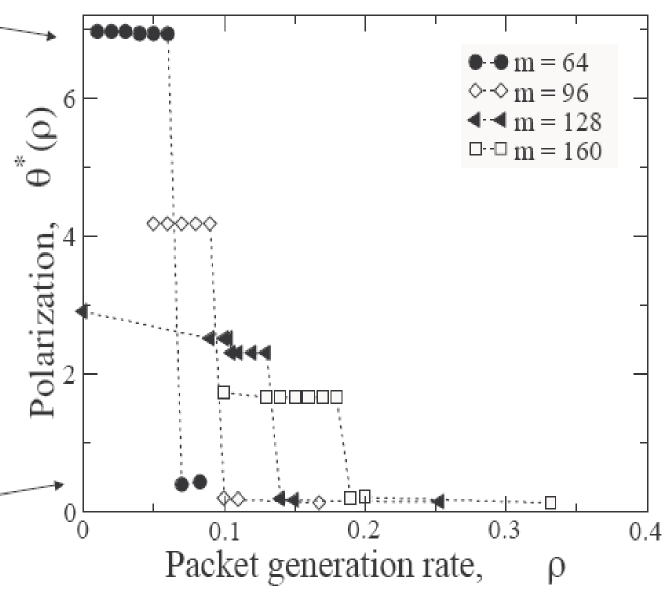


star

random



Polarization



$\langle \beta \rangle$: Average node betweenness

$$\theta = \frac{\max_{i \in N} \beta_i - \langle \beta \rangle}{\langle \beta \rangle}$$

Optimization in complex network, (Ferrere, 2006)

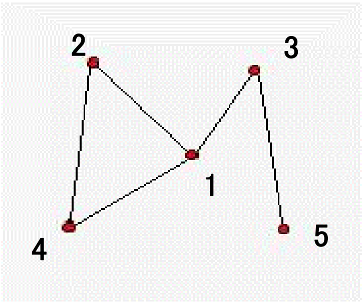
Adjacency Matrix of Graph

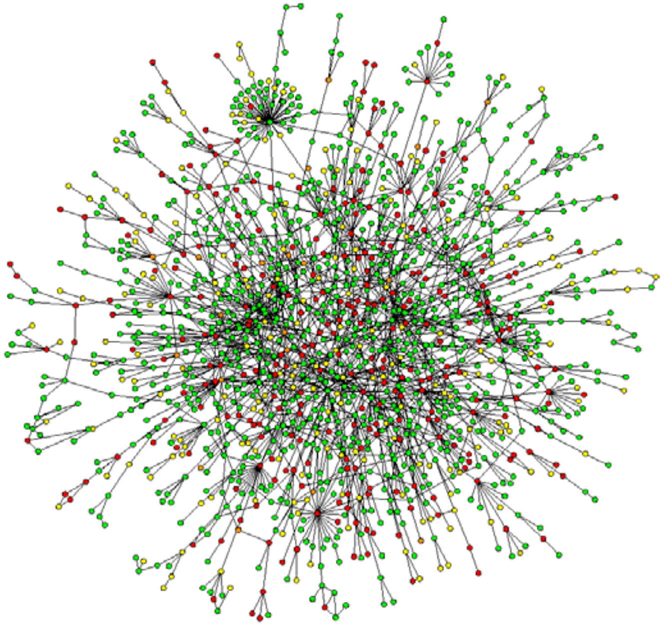
- The cording of the adjacency matrix: A=(a_{ij})
  - Node i and node j is connected : a_{ij}= 1
  - Node i and node j is not connected : a_{ij}= 0

Adjacency matrix

Example :

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

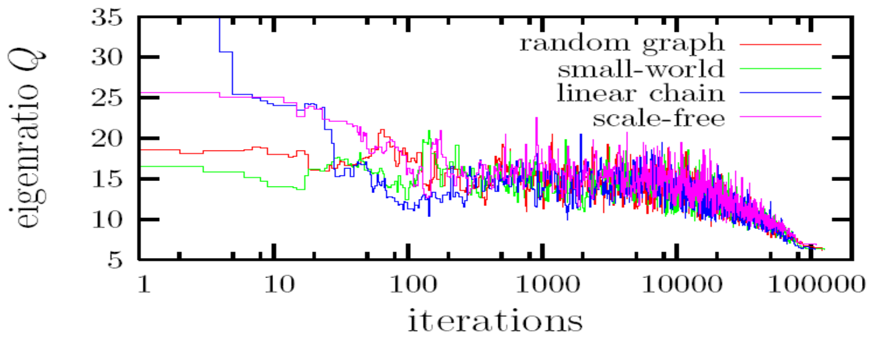


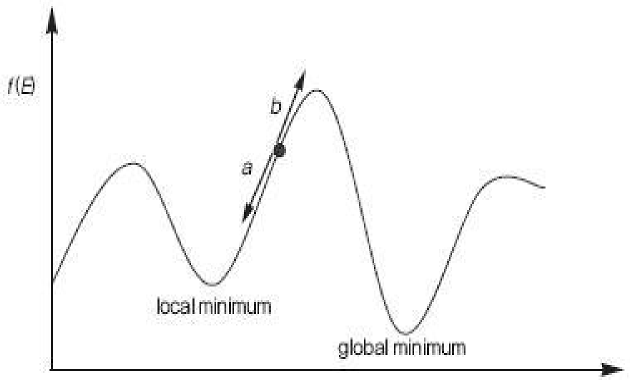


Stochastic Optimization

Simulated annealing

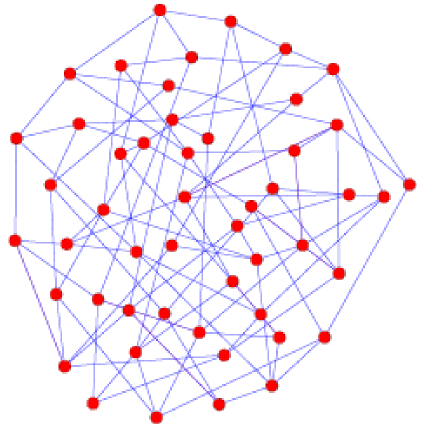
- Probabilistic algorithm for the optimization problem
- Rewiring trials - Rewiring a randomly selected link
- Fitness function to be optimized: Q
    - if $\delta Q = Q_{final} - Q_{initial} < 0$
      accept rewiring





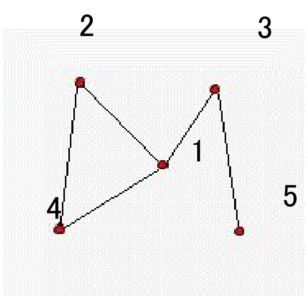
Optimized network

N = 50, and <k> = 4



The Fitness Function (1)
Link Density

- Design cost: the link density : α $$\alpha = \frac{1}{{}_nC_2} \sum_{i<j} a_{ij}$$
  - Maximum possible links of the network with n nodes: ${}_nC_2$
  - The number of links $$\sum_{i<j} a_{ij}$$



	1	2	3	4	5
1	0	1	1	1	0
2	1	0	0	1	0
3	1	0	0	0	1
4	1	1	0	0	0
5	0	0	1	0	0

$$\sum_{i<j} a_{ij} = 5$$

The Fitness Function (2) Congestion Index

- Congestion measure : $\lambda(\rho)$
  - Packet generation probability on certain node : $\frac{\rho}{n-1}$

    ρ : packet generation rate

  - Quantity of packet input on k node : $\frac{\rho}{n-1} \times \beta_k$
  - Quantity of packet output : 1 β_k : betweenness at k node
  - Queue length average on k node : $\frac{\rho \frac{\beta_k}{n-1}}{1-\rho \frac{\beta_k}{n-1}}$ Little's law

Congestion measure

- Total queue length on the network:

$$\lambda(\rho) = \sum_{k \in N} \frac{\rho \frac{\beta_k}{n-1}}{1-\rho \frac{\beta_k}{n-1}}$$

The Weighted Fitness Function

- Link density: α
- Congestion function: $\lambda(\rho)$
  - ρ : Packet generation rate
- Weight : ω $0 \le \omega \le 1$

The weighted object function to be minimized: $E(\omega, \rho)$

$$0 \le \omega \le 1$$

$$E(\omega, \rho) = \omega\lambda(\rho) + (1-\omega)\alpha$$

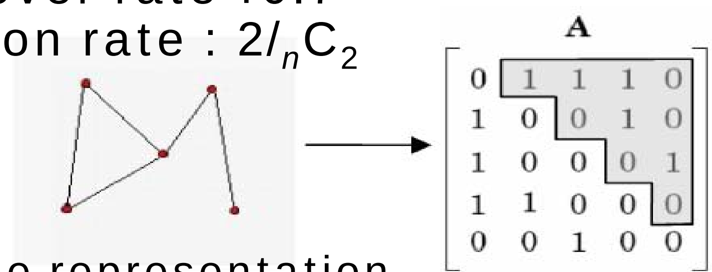
$$\lambda(\rho) = \sum_{k \in N} \frac{\rho \frac{\beta'_k}{n-1}}{1 - \rho \frac{\beta'_k}{n-1}} \qquad \alpha = \frac{1}{{}_nC_2} \sum_{i<j} a_{ij}$$

Generic Algorithm

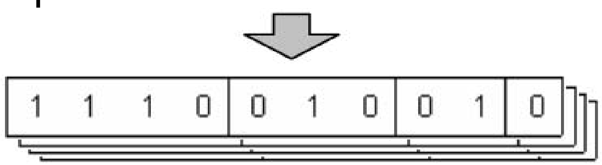
MGG Model(Minimal Generation Gap)

Crossover rate : 0.7

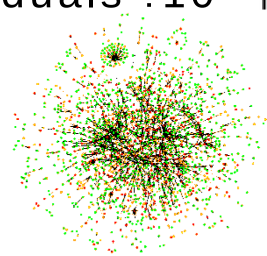
Mutation rate : $2/{}_nC_2$

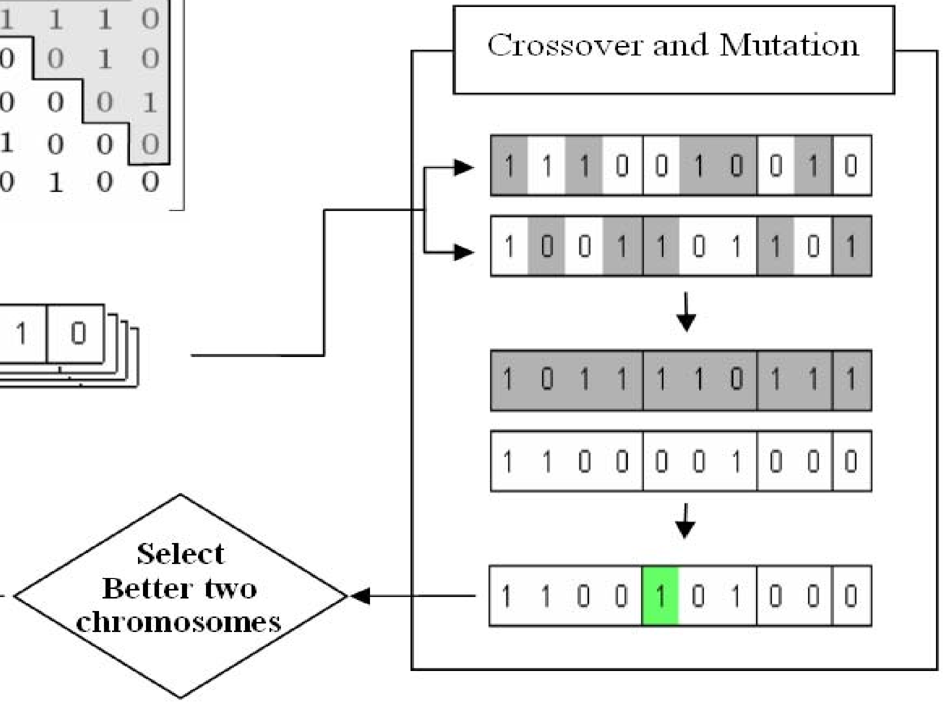


Generic code representation



Number of Individuals : 10 (networks)





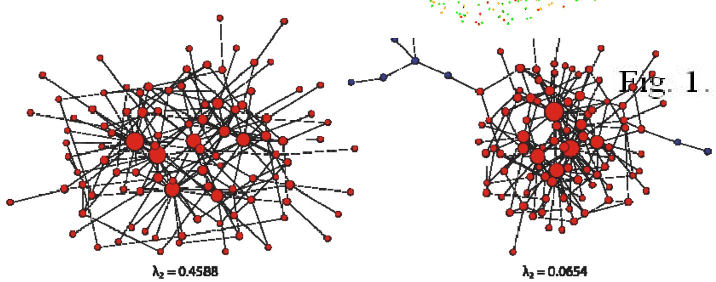
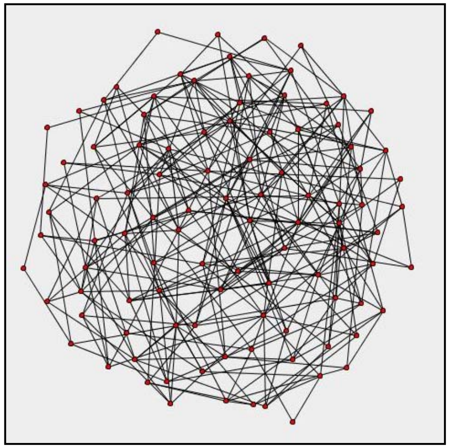


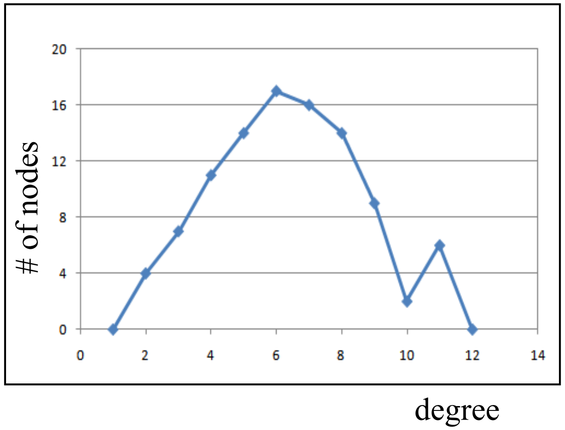
Fig. 1. The basic scheme of the minimization algorithm.

An Initial Network

Initial: Random network

- A fixed number of nodes : 100
- Links creation
  - Poisson distribution
  - 7 link per each node





The degree distribution

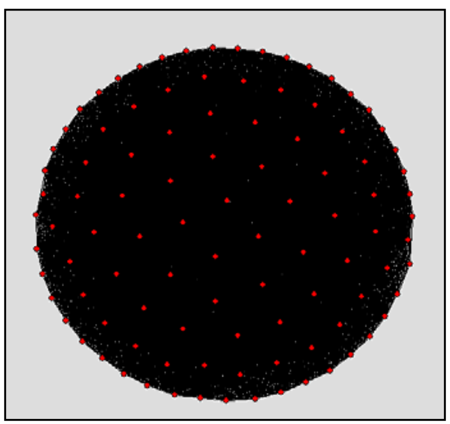
Optimized Network (1)

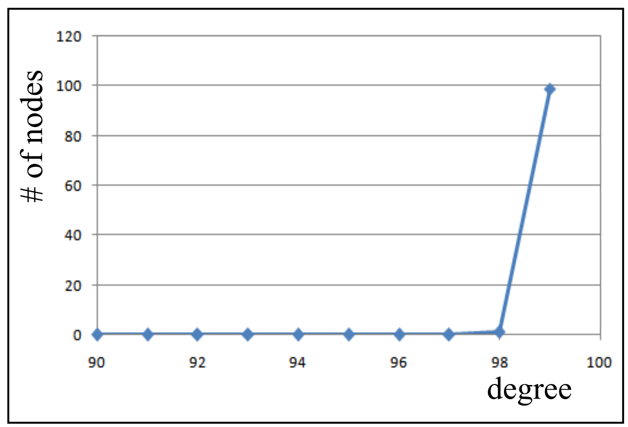
$$E(\omega,\rho)=\omega\lambda(\rho)+(1-\omega)\alpha$$

ω = 1 : Optimizing only congestion function

(packet generation rate: ρ=0.3)

- Optimal network: Complete network
  - Average link per node : 99.9
  - Link density : 0.9999 (4949/4950)
  - Congestion function value : 0→ no congestion





The degree distribution

Optimized Network (2)

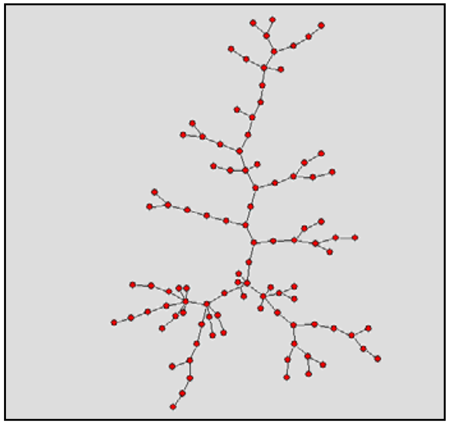
$$E(\omega,\rho)=\omega\lambda(\rho)+(1-\omega)\alpha$$

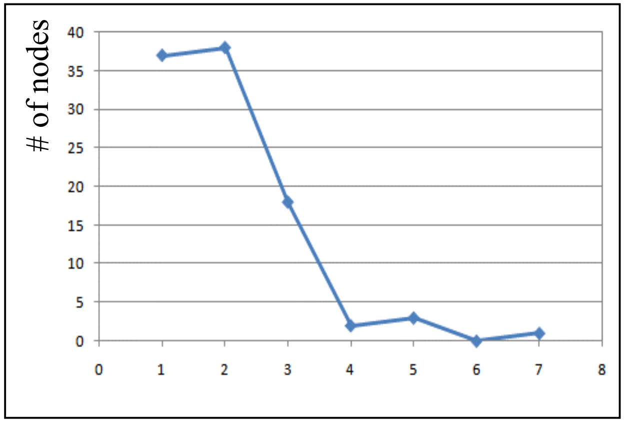
- ω = 0 : Minimizing only link density

  (packet generation rate: ρ=0.3)

  - Optimal network: Tree-like network
    - Average link :1.98
    - Link density :0.02(99/4950)
    - Congestion index :0.027

  → Tree structure has the smallest links





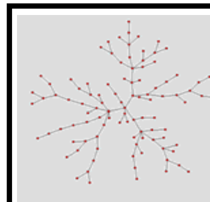
Optimized Networks (3)

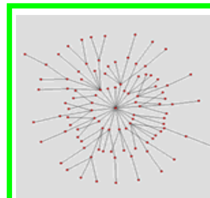
$$E(\omega,\rho)=\omega\lambda(\rho)+(1-\omega)\alpha \qquad 0\leq\omega\leq 1$$

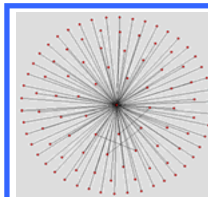
Low — link density — high

Packet generation: Low (↑) to High (↓)

ρ \ ω	0	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	1
0.05	LT	LT	LT	SH	SH	HT	SH	SH	MS	SS	SS	SS	C
0.1	LT	LT	SH	SH	MH	SH	SH	SS	SS	SS	SS	SS	C
0.2	LT	HT	HT	SH	SH	MS	MS	SS	SS	SS	R	R	C
0.3	LT	HT	SH	SH	MS	MS	SS	SS	SS	OS	R	R	C
0.4	LT	HT	SH	MS	MS	SS	OS	SS	SS	OS	R	R	C
0.5	LT	HT	HT	MS	MS	SS	OS	OS	OS	R	R	R	C
0.6	LT	MH	HT	MS	SS	OS	OS	OS	OS	R	R	C	C
0.7	LT	HT	MH	MS	SS	OS	OS	OS	OS	R	R	C	C
0.8	LT	MH	MH	MS	OS	OS	OS	OS	R	R	R	C	C
0.9	LT	MH	MS	MS	OS	OS	OS	R	R	R	R	C	C
0.95	LT	MH	MS	OS	OS	OS	OS	R	R	R	R	C	C

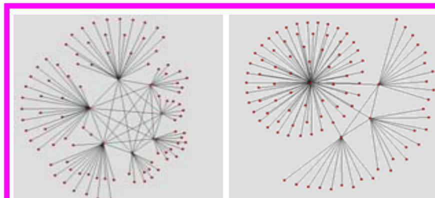

Line Like Tree

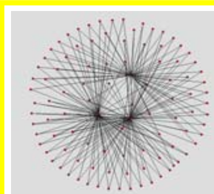

Single Hub

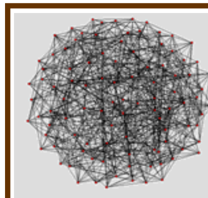

Single Star

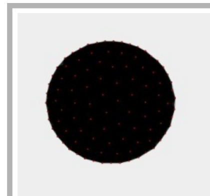

Hub & Tree


Multi Hub


Multi Star


Overlap Star


Random


Complete

Summary: Optimal Traffic Networks

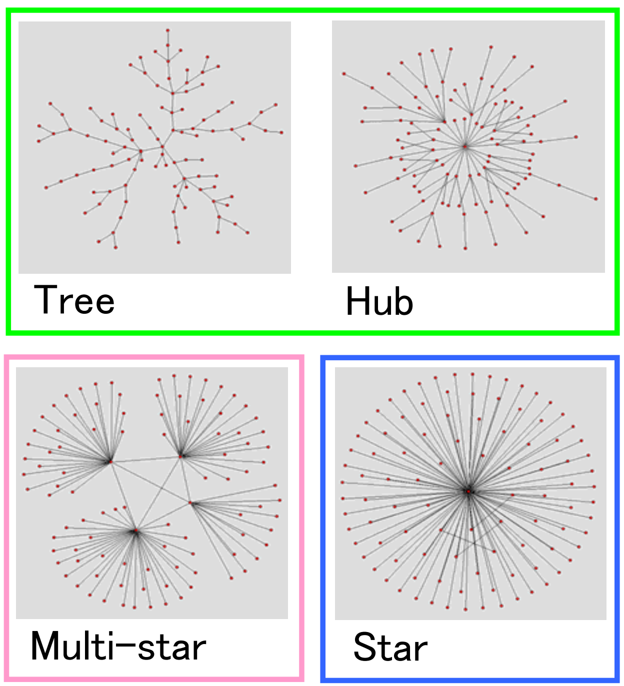
Phase 1: Tree → Hub → Star

:The link density increases slowly $\alpha \fallingdotseq 0.02 \rightarrow 0.15$

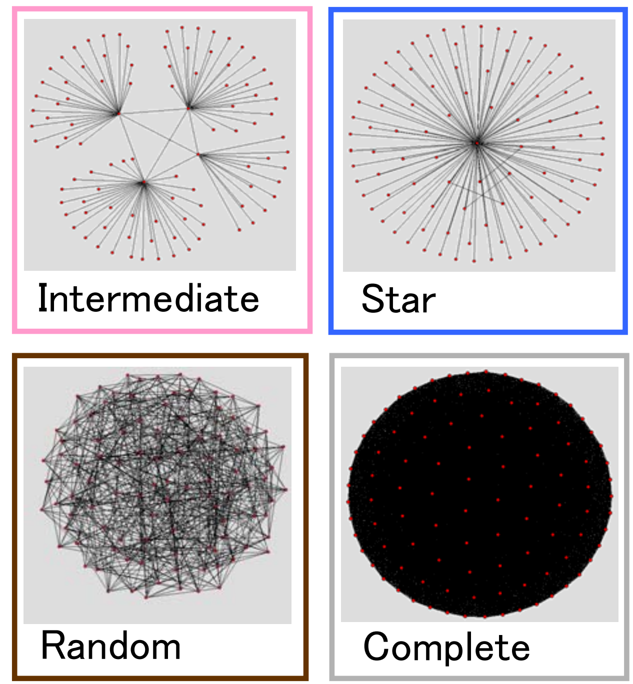
Phase 2: Star → Random → Complete

: Link density increases suddenly $\alpha = 0.15 \rightarrow \alpha \fallingdotseq 1$

Phase 1



Phase 2



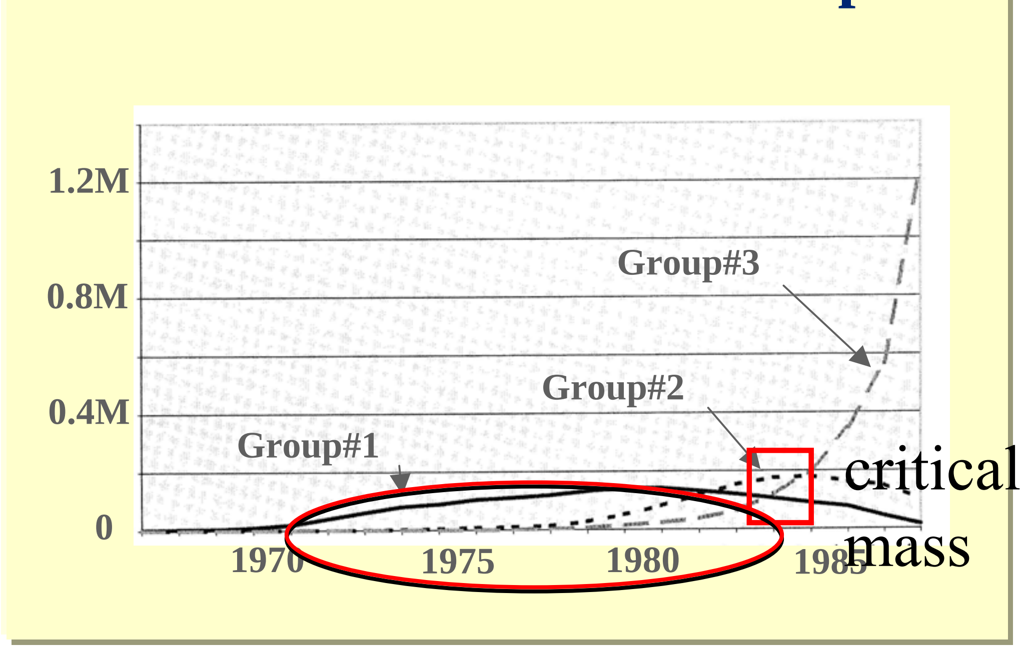
Emergence by Design: Evolutionary Optimization(2)

- Design of Communication Networks
  Tradeoff between congestion and design cost
- Diffusion of Innovation
- Consensus (synchronization) in Networked Systems

Diffusion of Innovation

- Why the markets occasionally accept innovations rather slowly compared with the superior technological advances of the innovation?
  "The slow pace of the fast change" (B. Chakravorti, 2003)

Competitive Innovations



Installed base of facsimile machine in North America (Rohlfs)

i-Phone

One SEG

Diffusion Models

- Concept of diffusion and contagion arises in many fields
  - Spread of infectious disease
  - Diffusion of innovations
  - Emergence of uncertainty in economic beliefs
  - Transmission of cultural fads

- Question 1: In what sense are these phenomena the same and how are they different?
- Question 2: What conditions trigger the decision to adopt something?

An Epidemic Diffusion Model (1)

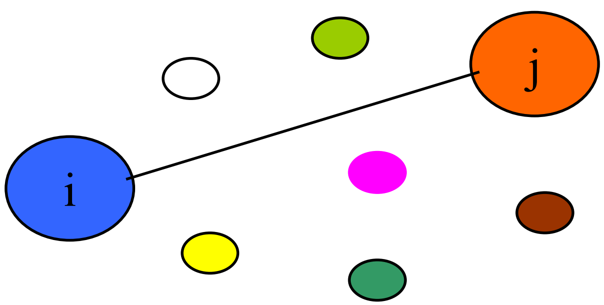
The SIR model

- Consider a fixed population of size N
- Each individual is in one of three states:
  - Susceptible (S), Infected (I), Removed (R)

$$S \xrightarrow[\beta]{} I \xrightarrow[\lambda]{} R$$

- Dynamic process: Mixing model

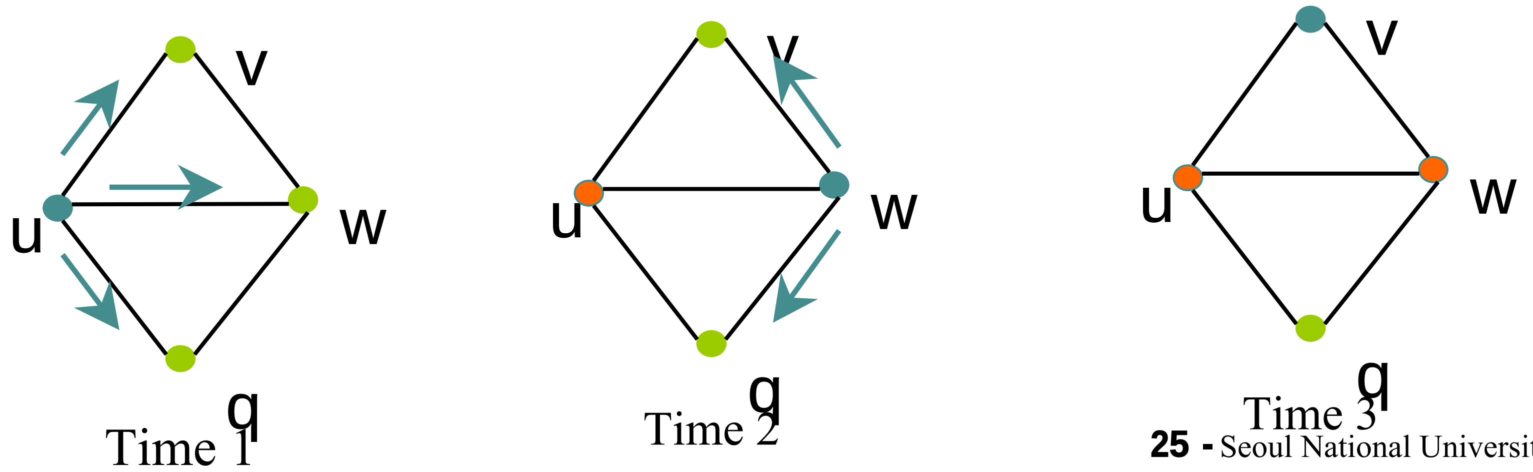
At each time step, each individual comes into contact with another individual chosen uniformly at random



An Epidemic Diffusion Model (2)

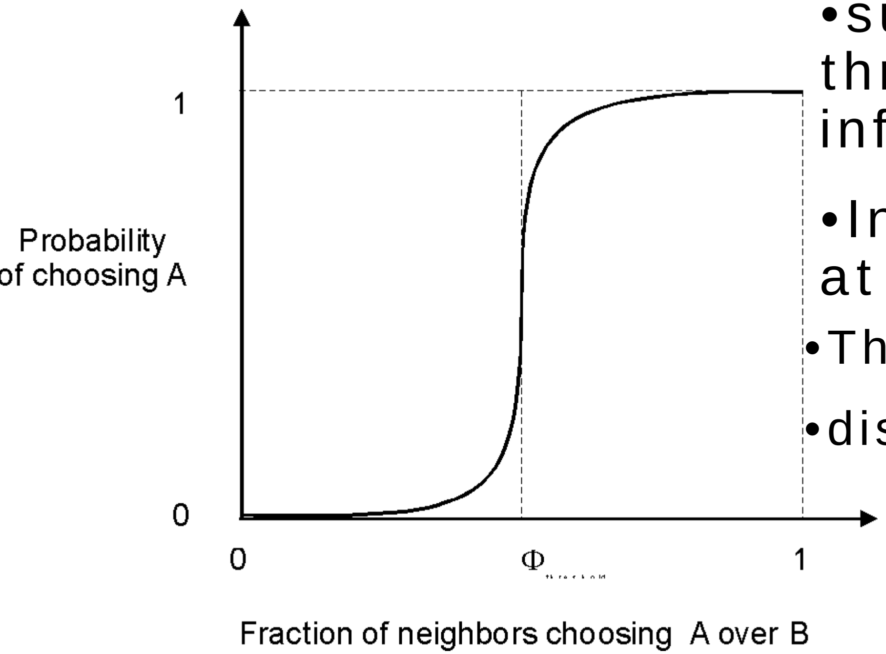
- Each node may be in the following states
  - Susceptible (S) (unaware, also inactive, non-adopter)
  - Infected (I) (aware, also active, informed, adopter)
  - Removed (R) (lose interest or forget)

- Infection rate β : probability of getting infected by a neighbor per unit time
- Immunization rate γ: probability of a node getting recovered per unit time



Universal Property of Diffusion Models

- Global Infection only occur after a threshold (critical mass)
- Many models on epidemic spreads, information cascades, fads, have the same threshold property



- susceptible become infected through their contacts with infected individuals at a rate β
- Infected agents are removed at rate γ
- There is a threshold above which the
- diseases spread through the population

$$\frac{\beta}{\gamma} = \lambda_c$$

λ_c Critical mass: threshold property in social dynamics

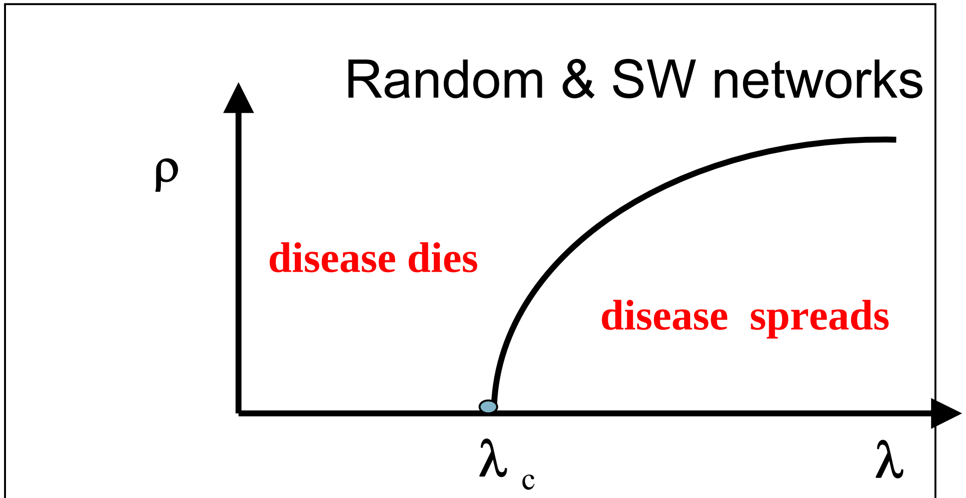
- The network topology affects critical mass

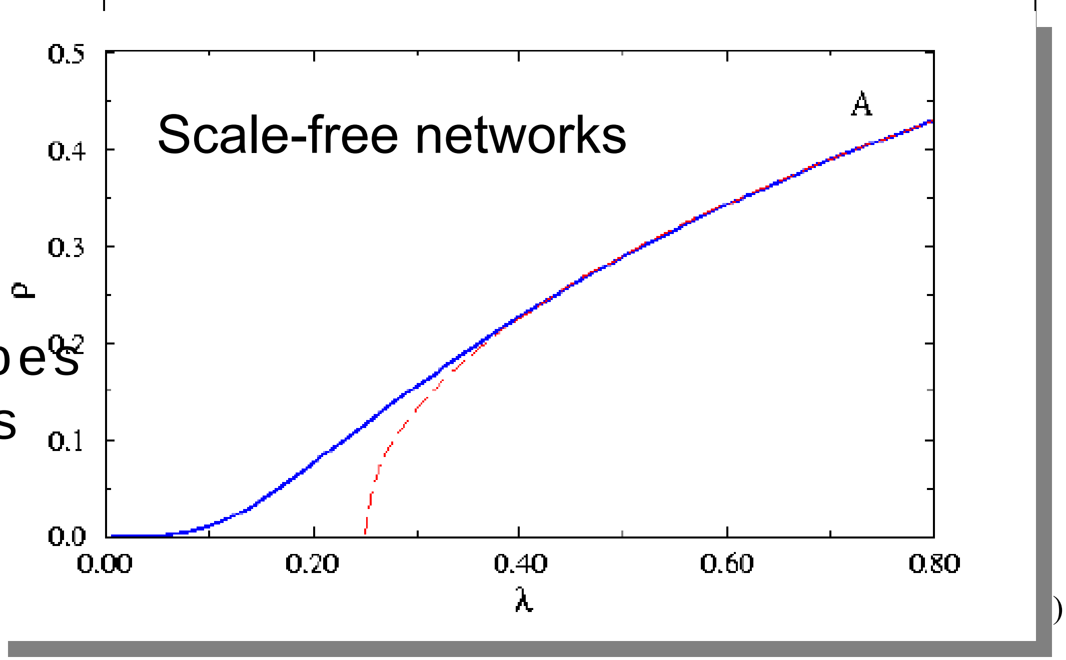
Network Topology & Critical Mass

The critical mass is given at

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

<k>: average degree of node

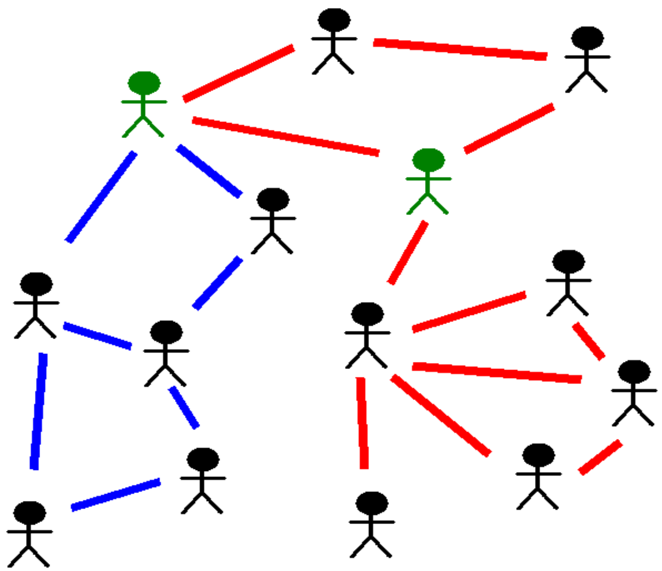




Scale-free network does not have critical mass

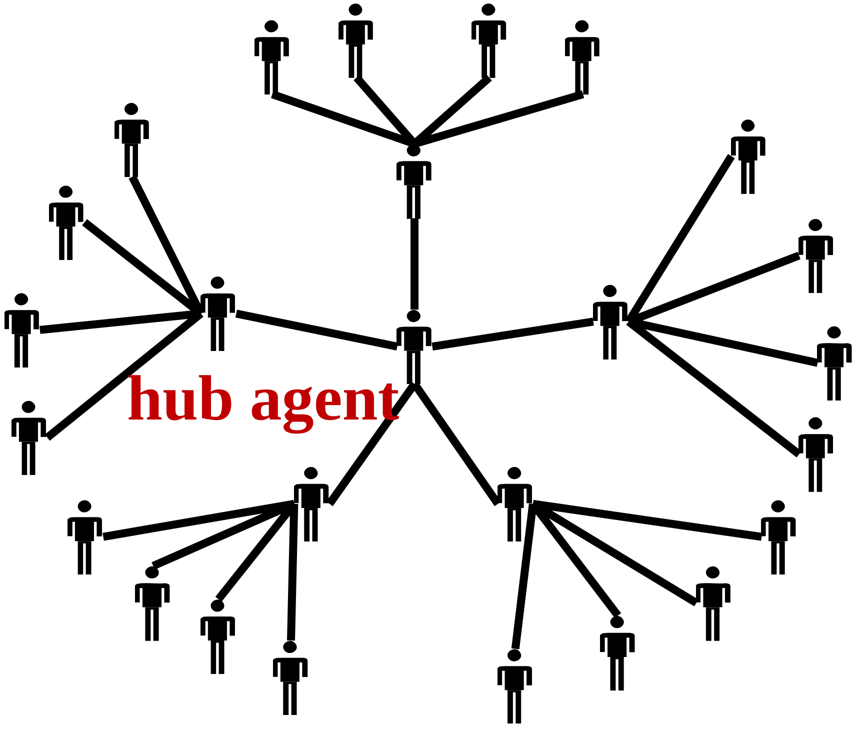
Social Influence Networks

Peer influence creates consensus
within *small* social groups



Local networks

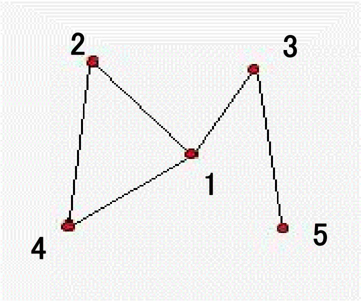
Impact of opinion leaders may be large



Scale-free networks

Dominant Eigenvalue of Adjacency Matrix

Example:



Adjacency matrix: symmetric

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Eigenvalue of symmetric matrix $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

$\lambda_1(A)$ is the largest eigenvalue of the adjacency matrix A

- The epidemic threshold is $\lambda_c = \beta/\gamma = 1/\lambda_1(A)$

Convergence of Diffusion Process

- The expected state of the system at time t is given by
- As t → ∞ $\overline{\mathbf{v}^{t}} = \left(p\mathbf{A} + (1-q)\mathbf{I}\right)\overline{\mathbf{v}^{t-1}}$
  - if $\lambda_1\left(p\mathbf{A} + (1-q)\mathbf{I}\right) < 1 \Leftrightarrow \lambda_1\left(\mathbf{A}\right) < \gamma/\beta$ then $\overline{\mathbf{v}^{t}} \to 0$
    - the probability that all copies die converges to 1
  - if $\lambda_1\left(p\mathbf{A} + (1-q)\mathbf{I}\right) = 1 \Leftrightarrow \lambda_1\left(\mathbf{A}\right) = \gamma/\beta$ then $\overline{\mathbf{v}^{t}} \to \mathbf{c}$
    - the probability that all copies die converges to 1
  - if $\lambda_1\left(p\mathbf{A} + (1-q)\mathbf{I}\right) > 1 \Leftrightarrow \lambda_1\left(\mathbf{A}\right) = \gamma/\beta$ then $\overline{\mathbf{v}^{t}} \to \infty$
    - the probability that all copies die converges to a constant < 1

$\lambda_1(A)$ is the largest eigenvalue of the adjacency matrix A

The Largest Eigenvalue

$$\lambda_{\max} = \max \frac{x^{\mathsf{T}} A x}{\|x\|^2}$$

$$d_{\text{average}} \le \lambda_{\max} \le d_{\max}$$

average degree

maximum degree

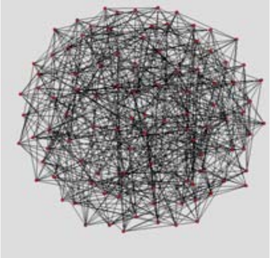
$$\sqrt{d_{\max}} \le \lambda_{\max} \le d_{\max}$$

If G is regular of degree d, then $\lambda_{\max} = d$.

An Eigenvalue Point of View

$\lambda_1(A)$: the largest eigenvalue of the adjacency matrix A

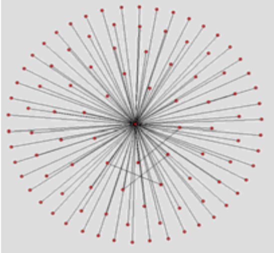
random



$\lambda_1 = pN = \langle k \rangle$

p: connection probability

star



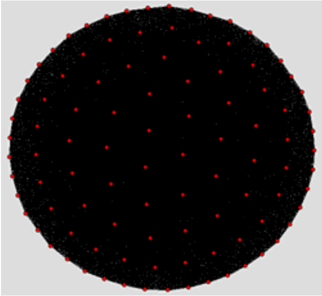
$\lambda_1 = \sqrt{N-1}$

Scale free

Multi Hub

$\lambda_1 \cong N^{1/4}$

complete



$\lambda_1 = N - 1$

Object 1: Minimizing spread of diffusion

Object 2: Maximizing spread of diffusion

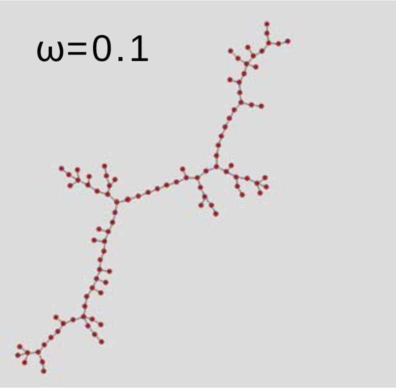
Minimizing Diffusion

- Object function 1: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

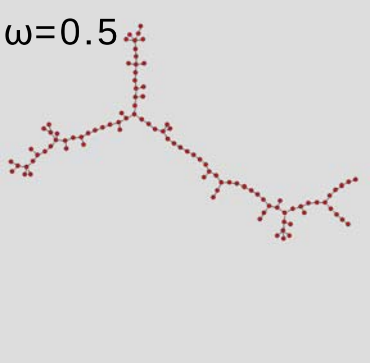
  Minimize the largest eigenvalue

  $$F = \omega\lambda_1 + (1-\omega) < k >$$

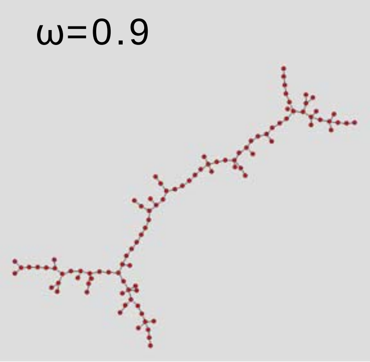
  <k>: average degree



λ_1=2.38 <k>=1.98
λ_n=-2.38



λ_1=2.38 <k>=1.98
λ_n=-2.38



λ_1=2.48 <k>=1.98
λ_n=-2.48

Line structure is optimal for minimizing diffusion

Maximizing Diffusion

- Object function 2:

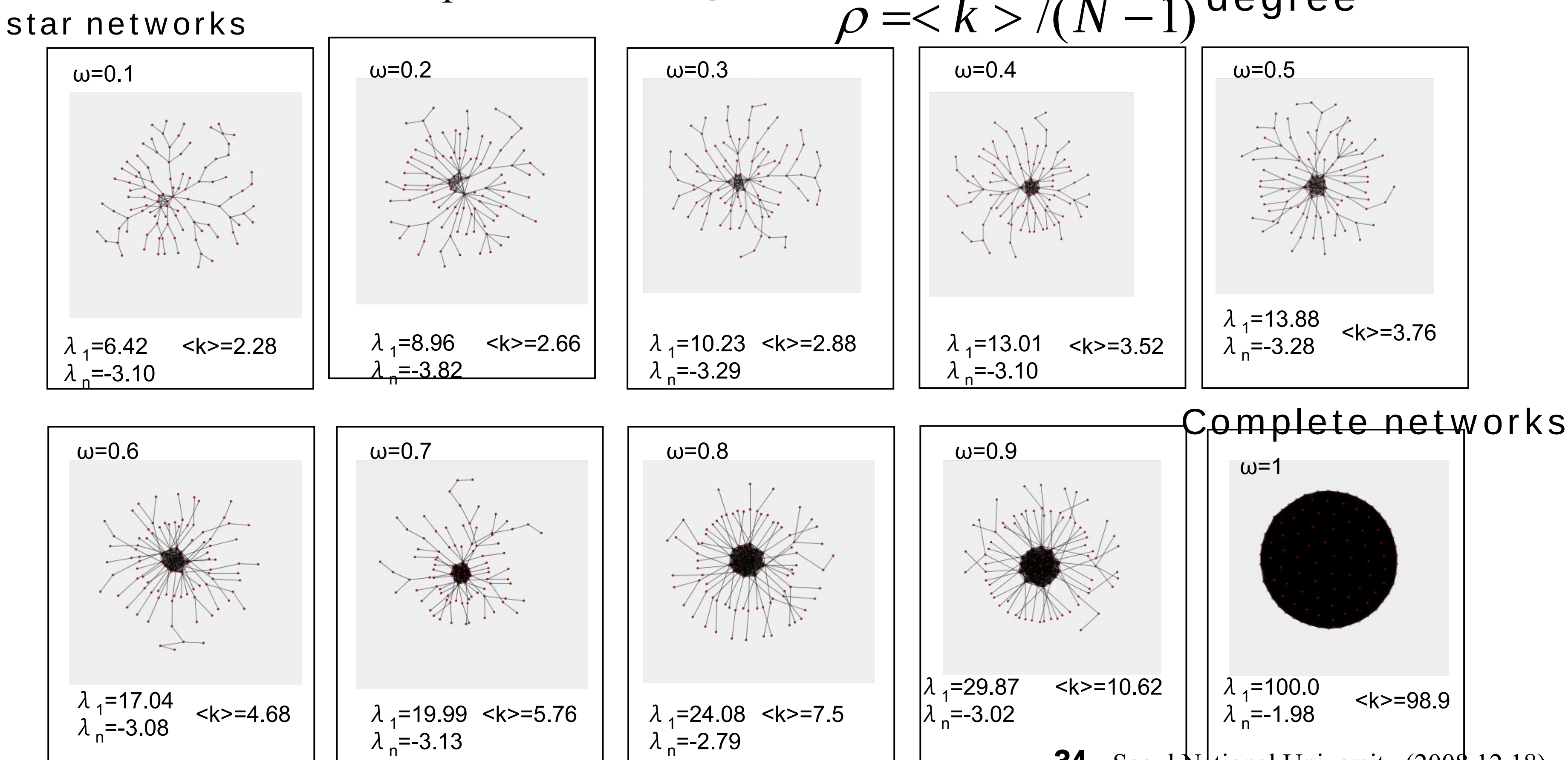
Maximize the largest eigenvalue

$$F = \omega / \lambda_1 + (1-\omega)\rho$$

ρ link density

$\rho = <k> / (N-1)$

<k>: average degree



Core dense network is optimal for maximizing diffusion

Emergence by Design: Evolutionary Optimization (3)

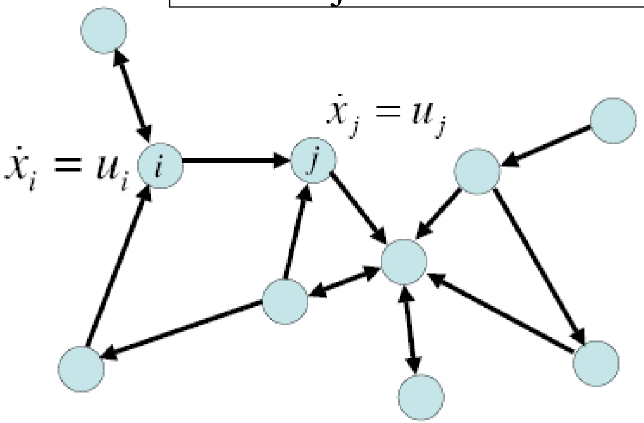
- Design of Communication Networks
  Tradeoff between congestion and design cost
- Diffusion of Innovation
- Consensus (synchronization) in Networked Systems

Consensus Problems

"Consensus" means to reach an agreement regarding a certain quantity of interest that depends on the state of all nodes (subsystems).

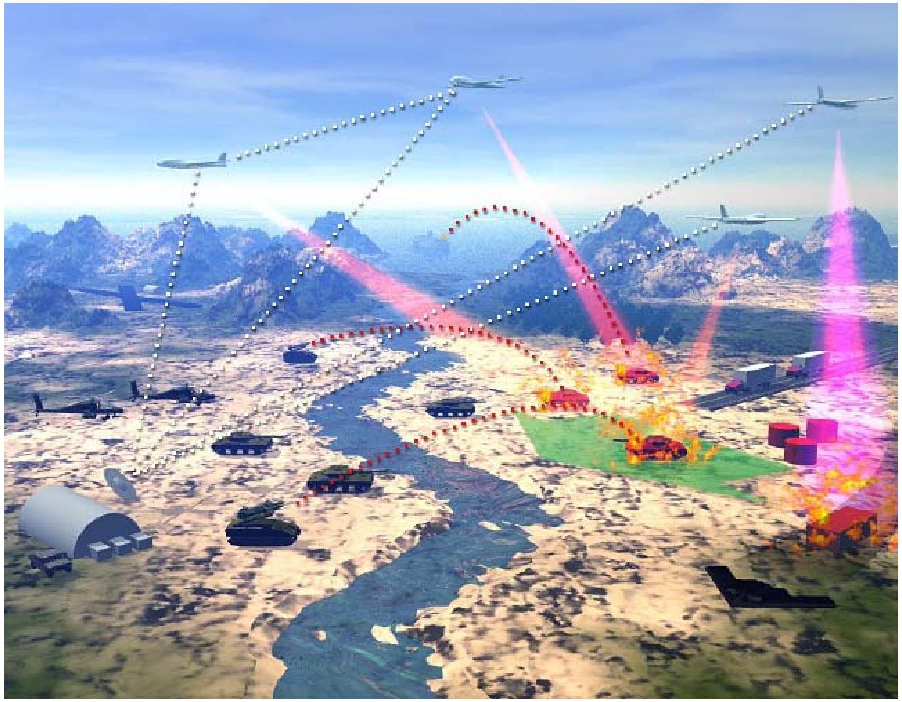
More specific, a consensus algorithm is a decentralized rule that results in the convergence of the states of all network nodes to a common value.

$$x_i = x_j = \ldots = x_{consensus}$$



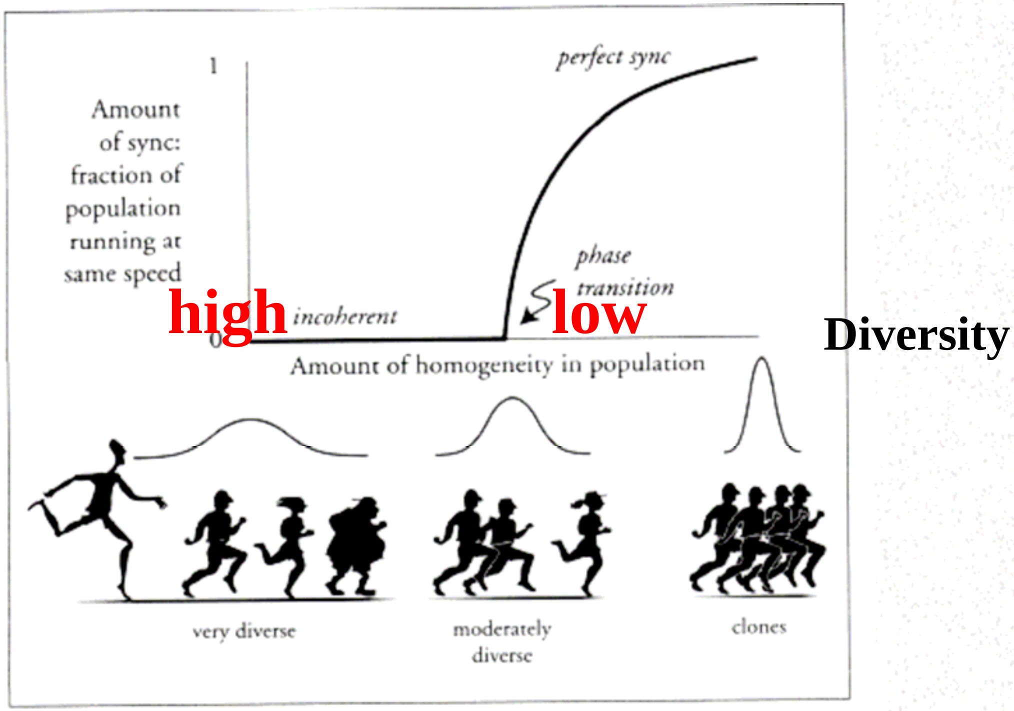
Consensus Problems in Engineering

A position reached by a group as a whole



Battle space management scenario illustrating distributed command and control between heterogeneous air and ground assets

Synchronization



Synchronization: Prevalent appearance in physics and biology

Homogeneity is important for better synchronization

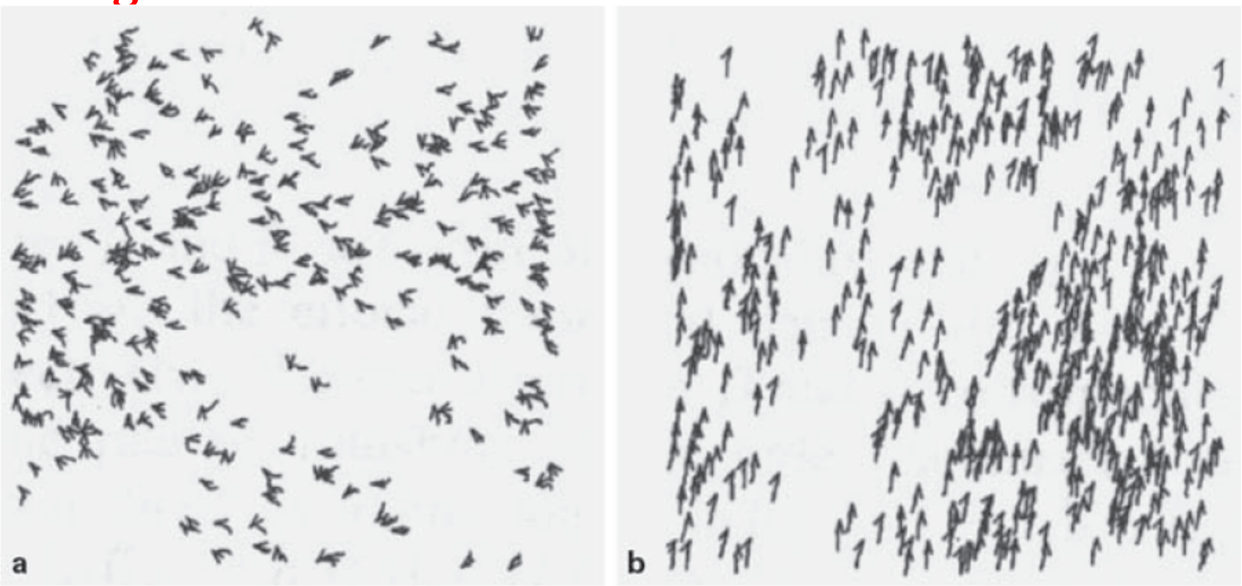
Consensus Problems & Synchronization

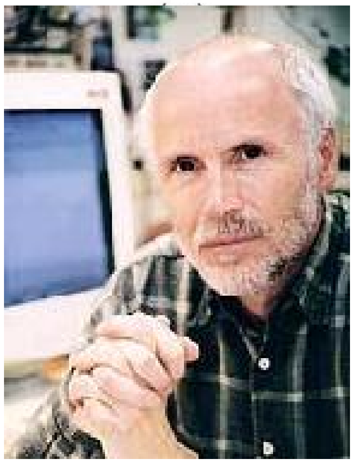
"Consensus has connections to problem in synchronization"

Model: every bird adjusts its velocity by adding to it a weighted average of the differences of its velocity with those of the other birds. That is, at time $t \in \mathbb{N}$, and for bird i,

$$v_i(t+1) - v_i(t) = \sum_{j=1}^{k} a_{ij}(v_j(t) - v_i(t)).$$

"Emergent behavior on flocks"

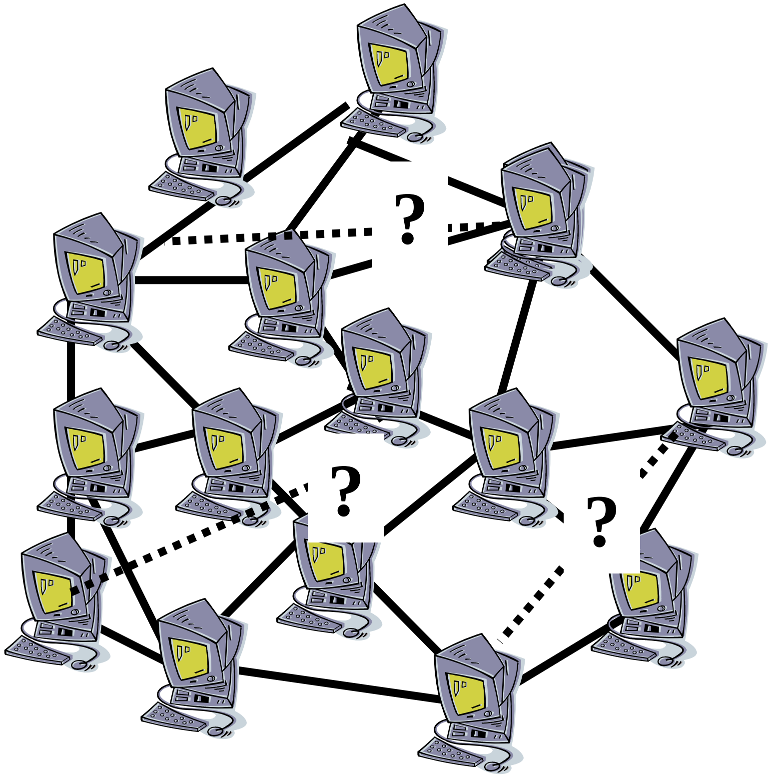




Vicsek T,.Phys Rev Letter (1995)

(2008.12.18)

Engineering Problems



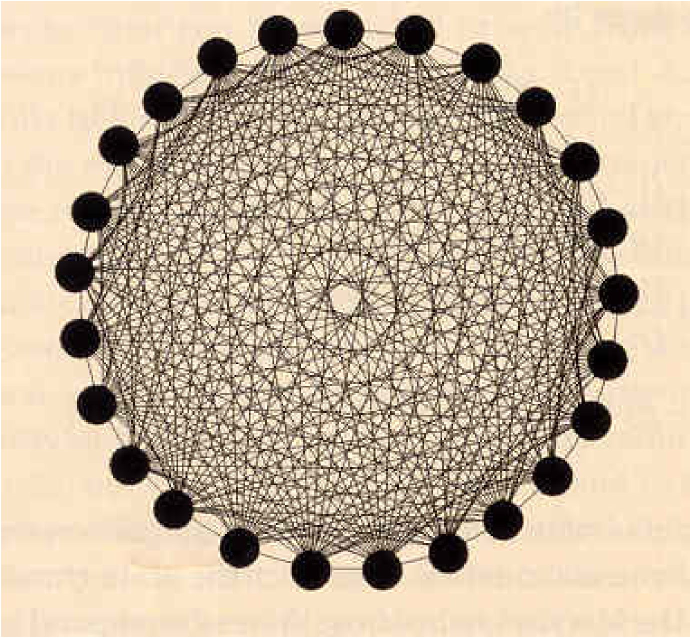
Question: How do we add some new links with better consensus?

Synchronization in Globally Connected Networks

Observation:

No matter how large the network is, a globally coupled network <u>will</u> synchronize if its coupling strength is sufficiently strong

Good – if synchronization is useful



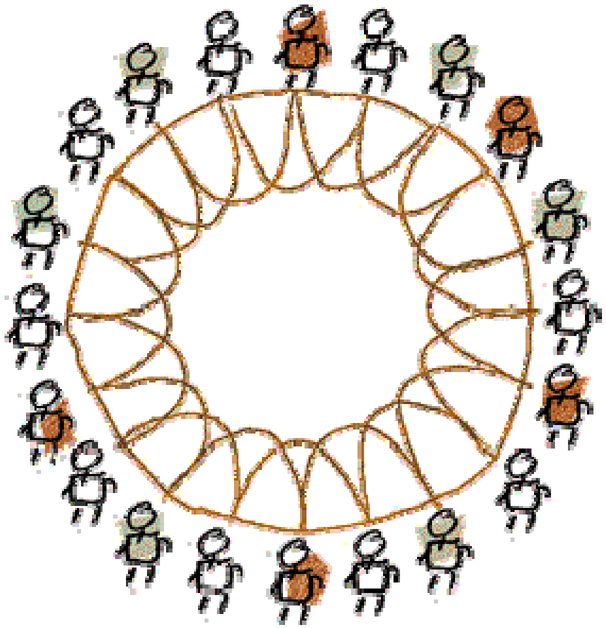
G. Ron Chen (2006)

Synchronization in Locally Connected Networks

Observation:

No matter how strong the coupling strength is, a locally coupled network <u>will</u> <u>not</u> synchronize if its size is sufficiently large

<u>Good</u> - if synchronization is harmful



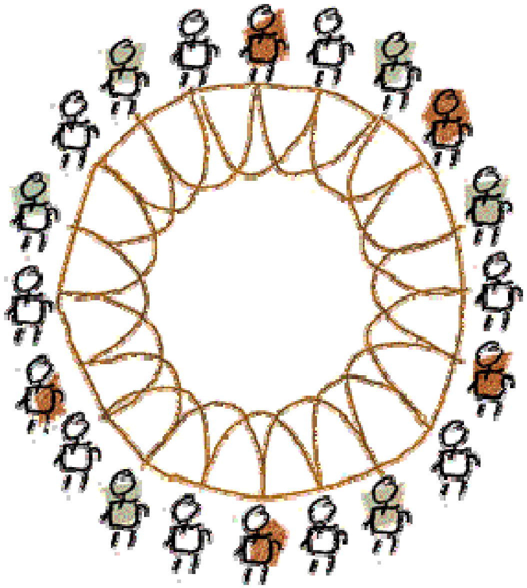
G. Ron Chen (2006)

Synchronization in Small-World Networks

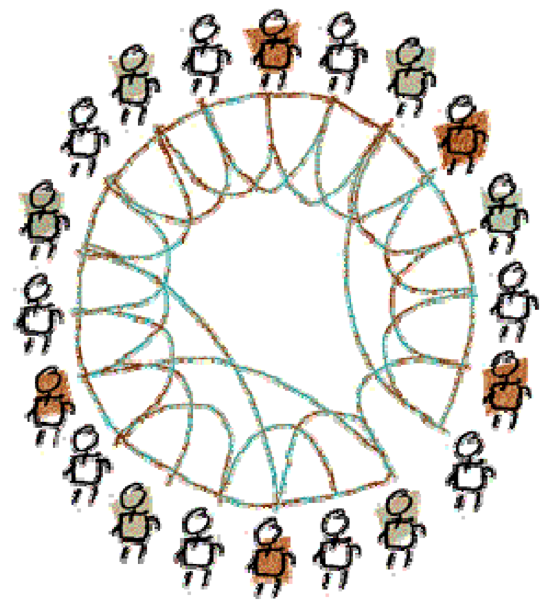
Start from a nearest neighbor coupled network

G. Ron Chen (2006)

small-world network



Add a link, with probability p, between a pair of nodes



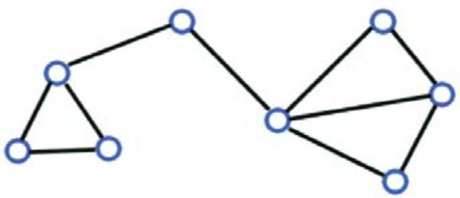
Good news: A small-world network is easy to synchronize!

X.F.Wang and G.R.Chen: Int. J. Bifurcation & Chaos (2001)

Synchronization & Network Topology

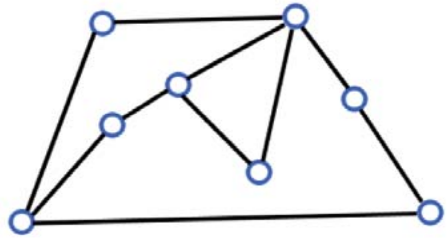
Connectivity of networks does matter for synchronization

Network A



$\lambda_2 = 0.238$

Network B



$\lambda_2 = 0.925$

Laplacian matrix

$$\begin{bmatrix} k_1 & & & \{0,-1\} \\ & k_2 & & \\ & & \ddots & \\ \{0,-1\} & & & k_n \end{bmatrix}$$

- **Laplacian matrix = Degree – Adjacency matrix**
- **$\lambda_1 = 0$ is always an eigenvalue of a Laplacian matrix**

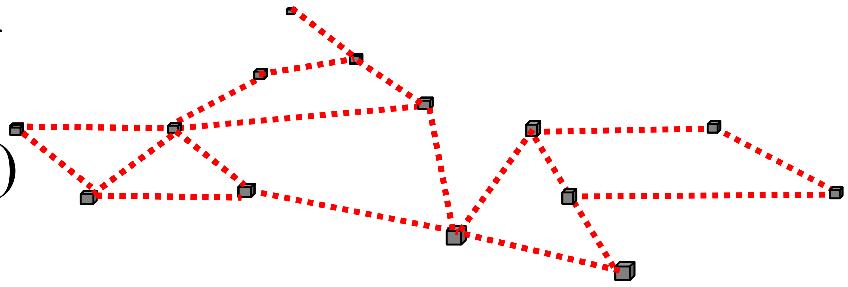
$$0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \le 2\Delta \qquad \Delta = \max_i d_i$$

λ_n / λ_2 : algebraic connectivity

: Smaller algebraic connectivity
: better consensus formation

Consensus Problems & Network Topology

The distributed consensus algorithm



$$x_i(t+1) = x_i(t) + \varepsilon \sum_{i \in N_i} w_{ij}(x_j(t) - x_i(t))$$

Convergence to the average of the initial values of all agents

$$x_1 = x_2 = \ldots = x_n = \sum_i x_i(0)/n$$

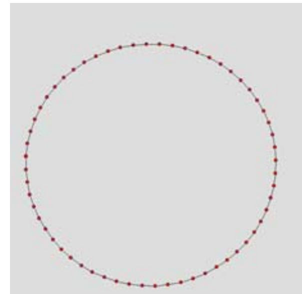
The weighted adjacency matrix $G=(w_{ij})$

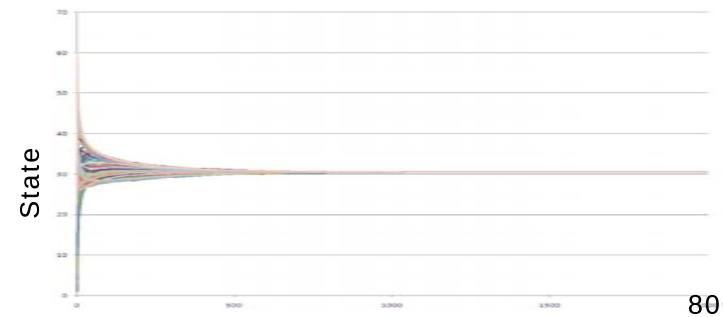
(i) Graph G is connected
(ii) G is balanced: symmetric graph

$$\sum_{i \neq j} w_{ij} = \sum_{j \neq i} w_{ji}$$

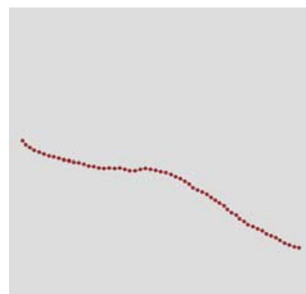
Convergence in Consensus Problems

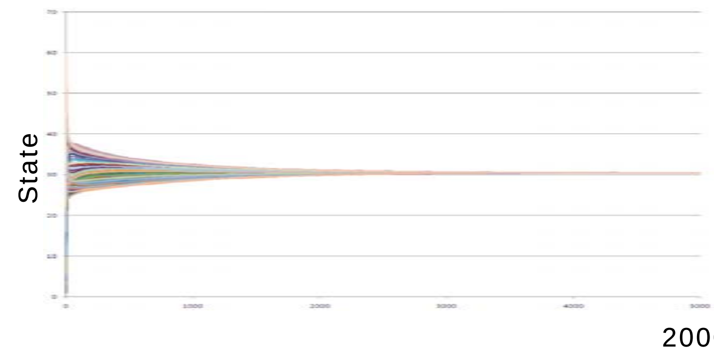
- circle
  - average link : 2
  - $\lambda_n / \lambda_2 = 365$
    - λ_2 : 0.01
    - λ_n : 4



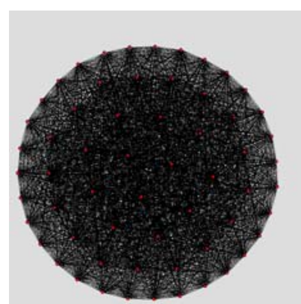


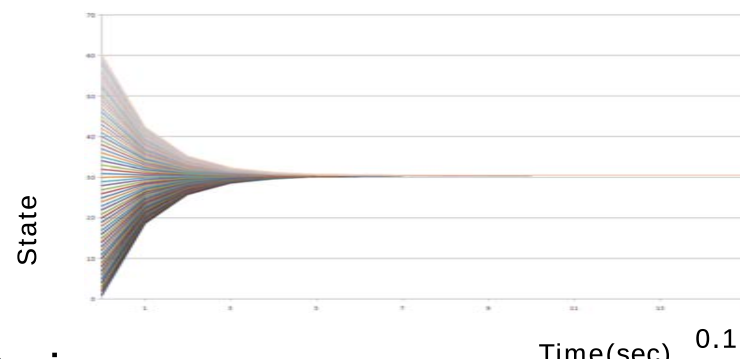
- line
  - average link : 2
  - $\lambda_n / \lambda_2 = 1458$
    - λ_2 : 0.003
    - λ_n : 4





- complete network
  - average link : 60
  - $\lambda_n / \lambda_2 = 1$
    - λ_2 : 60
    - λ_n : 60



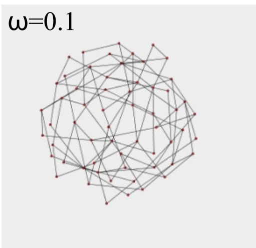


- Initial value of each agent : $x_i(o)=i$

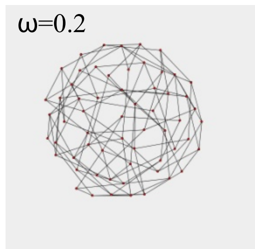
Optimized Networks (1)

$$E(\omega) = \omega \cdot \frac{\lambda_n}{\lambda_2} + (1-\omega)\cdot \alpha$$

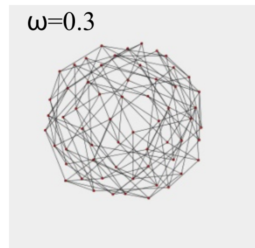
Ramanujam network



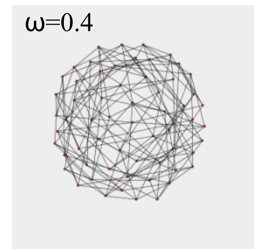
λ_2=0.7
λ_n=7.57
λ_n/λ_2=10.81
α=3.3
P=1.36



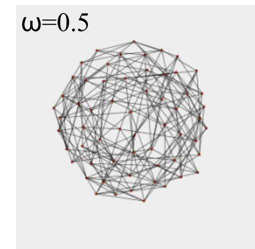
λ_2=1.07
λ_n=8.00
λ_n/λ_2=7.38
α=3.9
P=1.35



λ_2=1.65
λ_n=8.64
λ_n/λ_2=5.25
α=4.7
P=0.68

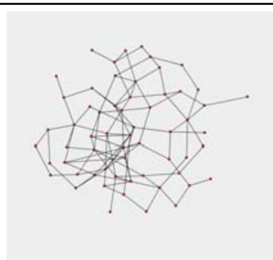


λ_2=2.16
λ_n=9.49
λ_n/λ_2=4.39
α=5.3
P=0.64

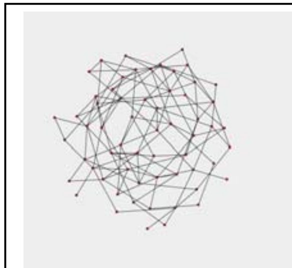


λ_2=2.83
λ_n=10.69
λ_n/λ_2=3.78
α=6.2
P=0.66

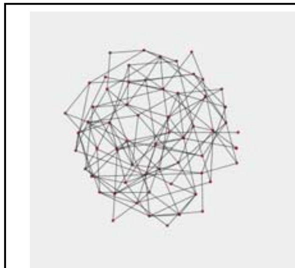
Random networks



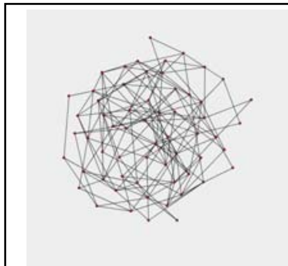
λ_2=0.27
λ_n=8.92
λ_n/λ_2=33.2
α=3.4
P=2.23



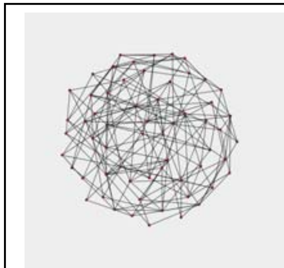
λ_2=0.48
λ_n=9.54
λ_n/λ_2=20.0
α=4.0
P=2.41



λ_2=0.77
λ_n=10.64
λ_n/λ_2=13.9
α=4.7
P=1.4



λ_2=0.96
λ_n=11.71
λ_n/λ_2=12.3
α=5.3
P=1.82

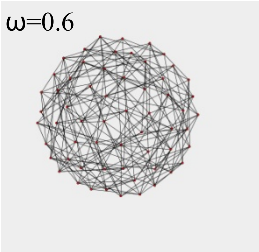


λ_2=1.73
λ_n=13.47
λ_n/λ_2=7.8
α=6.2
P=2.2

Optimized Networks (2)

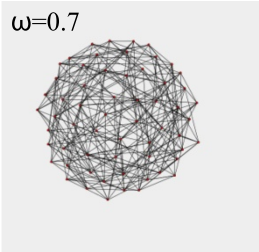
$$E(\omega) = \omega \cdot \frac{\lambda_n}{\lambda_2} + (1-\omega)\cdot \alpha$$

ω=0.6



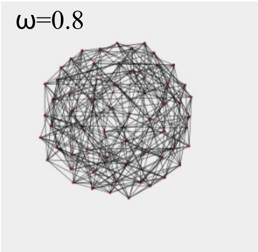
λ_2=3.4 α=6.9
λ_n=11.46 P=0.47
λ_n/λ_2=3.35

ω=0.7



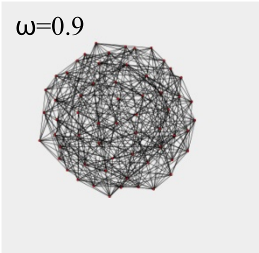
λ_2=4.64 α=8.3
λ_n=13.35 P=0.51
λ_n/λ_2=2.87

ω=0.8



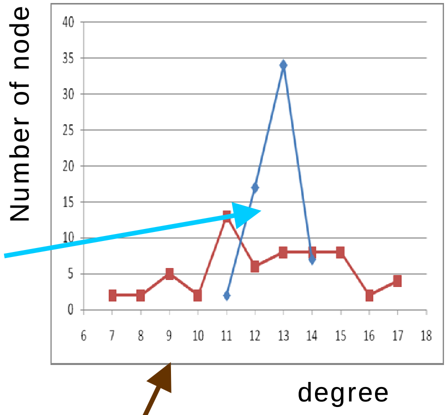
λ_2=6.17 α=10.2
λ_n=15.63 P=0.50
λ_n/λ_2=2.53

ω=0.9



λ_2=8.34 α=12.8
λ_n=18.57 P=0.31
λ_n/λ_2=2.23

Ramanujan network

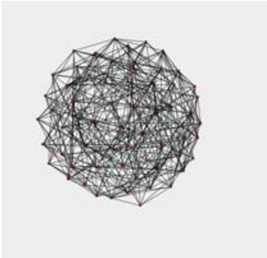


Random network

λ_2=1.58 α=6.9
λ_n=13.89 P=1.37
λ_n/λ_2=8.75

λ_2=2.08 α=8.3
λ_n=15.86 P=1.99
λ_n/λ_2=7.6

λ_2=3.87 α=10.2
λ_n=19.15 P=1.77
λ_n/λ_2=4.95



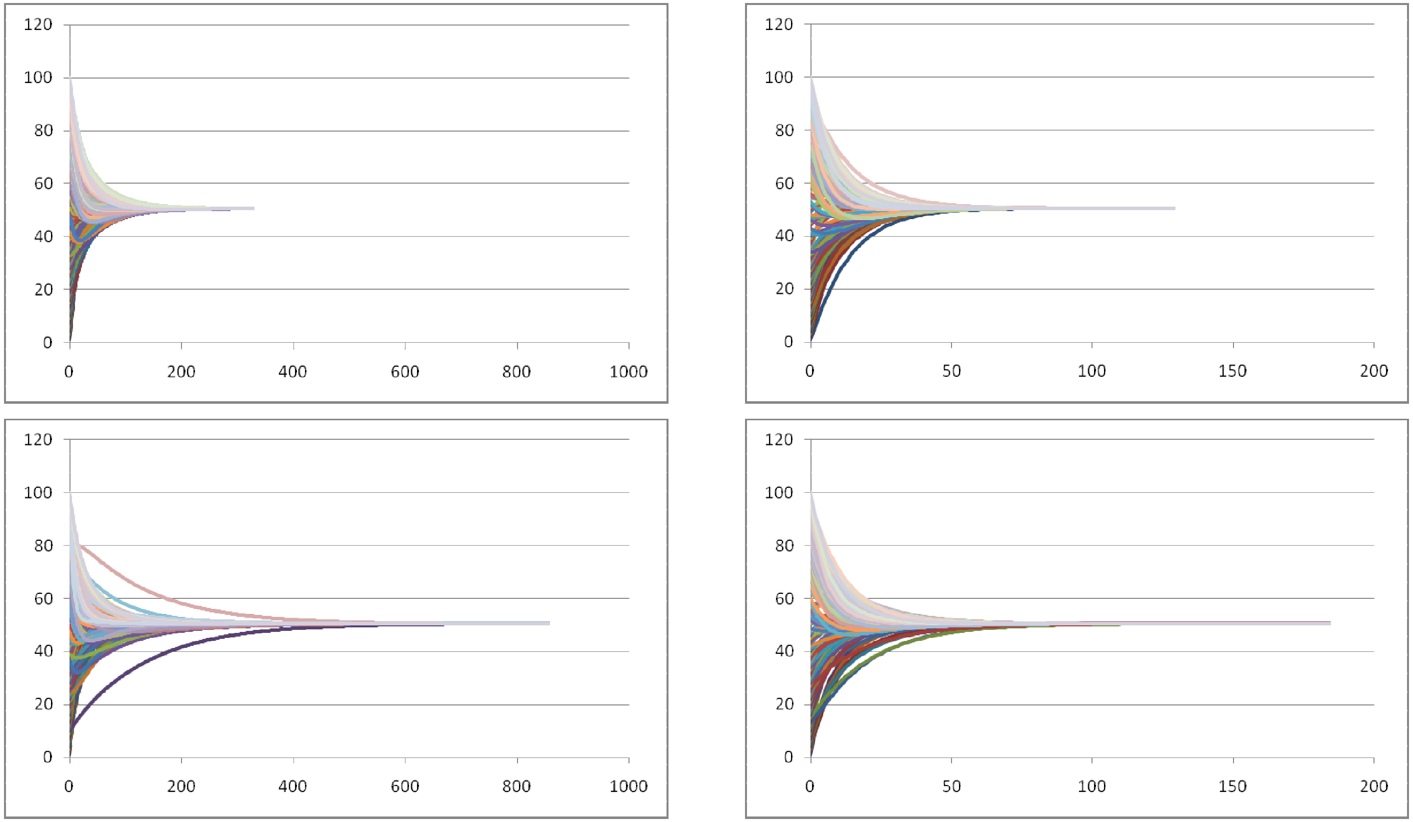
λ_2=5.53 α=12.4
λ_n=21.02 P=0.95
λ_n/λ_2=3.8

Comparison of Convergence Speed

- Initial value of each agent: $x_i(o)=i$

Optimal network

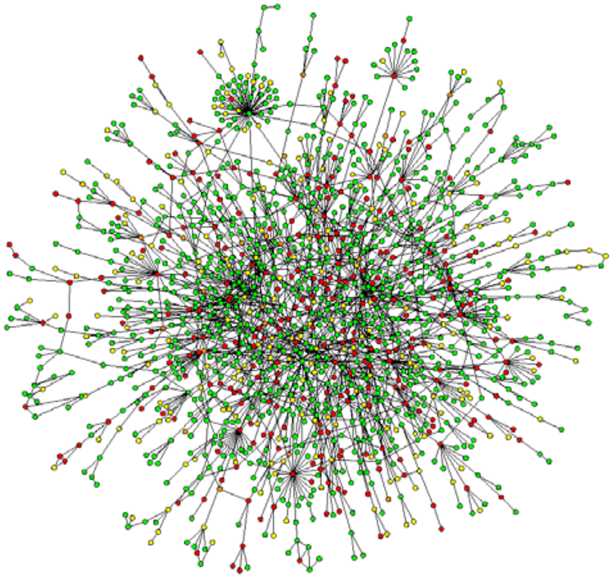
Random network



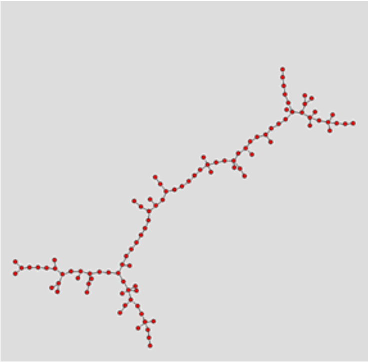
average link=5

average link= 11

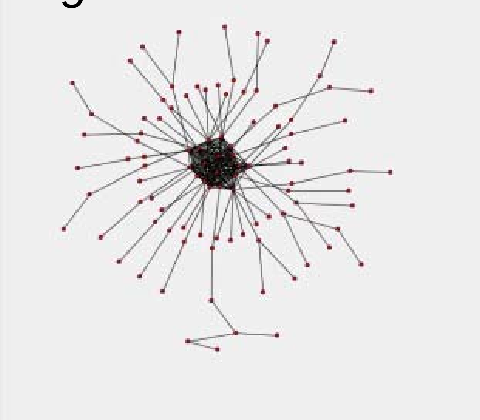
Multi-graph Topologies



- Minimizing spread
- *Minimize the spread of cascade failure or infective diseases*

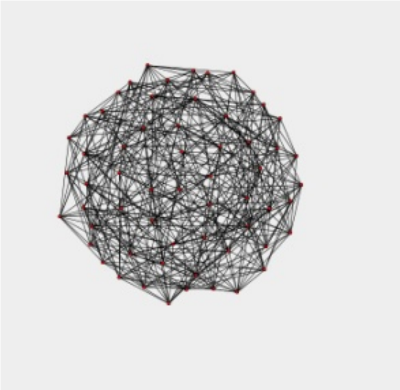


- Maximizing spread
- *Maximize the influence in voting campaign*



Synchronization,

Maximize the effect of coordinated behavior



Five Stages of Research

1) Observe: Gather data to demonstrate power law behavior in a system.
2) Interpret: Explain the import of this observation in the system context.
3) Model: Propose an underlying model for the observed behavior of the system.
4) Validate: Find data to validate (and if necessary specialize or modify) the model.
5) Design (Control): Design ways to control and modify the underlying behavior of the system based on the model.

Lots of open research problems in the design of complex systems

Conclusion

- Social systems involve a large-scale self-interested individual decisions that are main obstacles as well as driven forces for improving social systems.

- Social improvements that requires persuasion and consensus among us become very slow since most social influence networks are asymmetric

- *Evolutionary optimization is a powerful method for designing desirable social systems.*

Thank you for listening!!

Question Time