Complex Dynamic Networks: Architectures, Games, Components, Probability

John S. Baras

Institute for Systems Research
Department of Electrical and Computer Engineering
Fischell Department of Bioengineering
Applied Mathematics, Statistics and Scientific Computation Program
University of Maryland College Park

May 11, 2010
Presentation to
The Large Scale Networking (LSN)
Coordinating Group (CG) of the NITR

Acknowledgments

- **Joint work with:** Pedram Hovareshti, Tao Jiang, Ion Matei, Kiran Somasundaram, George Theodorakopoulos

- **Sponsors:** ARO (Wireless Network Security CIP URI, Robust MANET MURI), AFOSR (Distributed Learning and Information Dynamics MURI), ARL (CTA on C&N), NSF, DARPA (Dynamic Coalitions), Lockheed Martin, Telcordia

Taxonomy of Networked Systems

Infrastructure / Communication Networks

Internet / WWW
MANET
Sensor Nets
Robotic Nets
Hybrid Nets: Comm, Sensor, Robotic and Human Nets

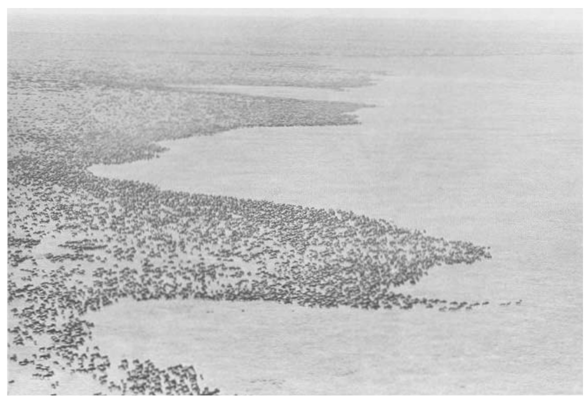
Social / Economic Networtks

Social Interactions
Collaboration
Social Filtering
Economic Alliances
Web-based social systems

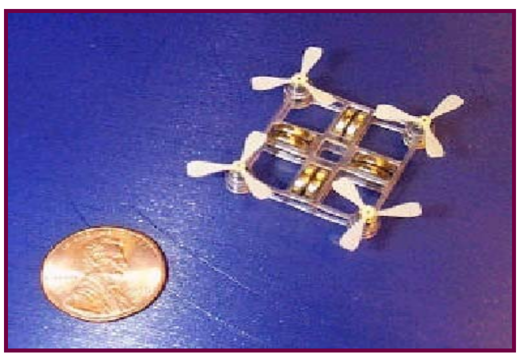
Biological Networks

Community
Epidemic
Cellular and Sub-cellular
Neural
Insects
Animal Flocks

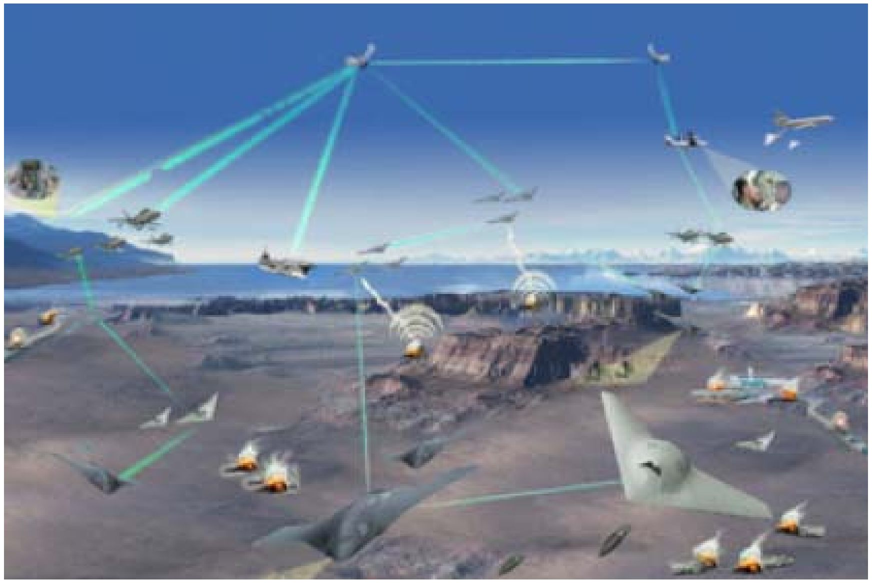
Biological Swarms

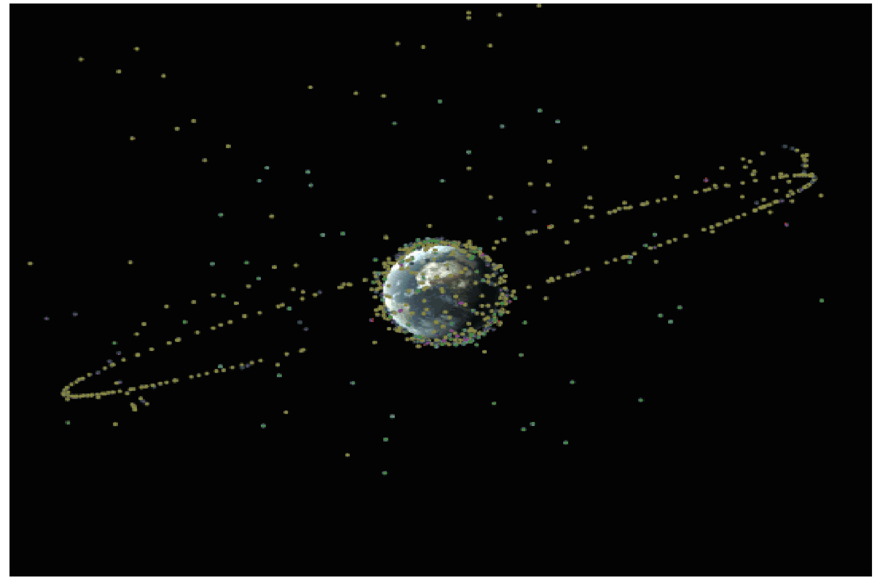


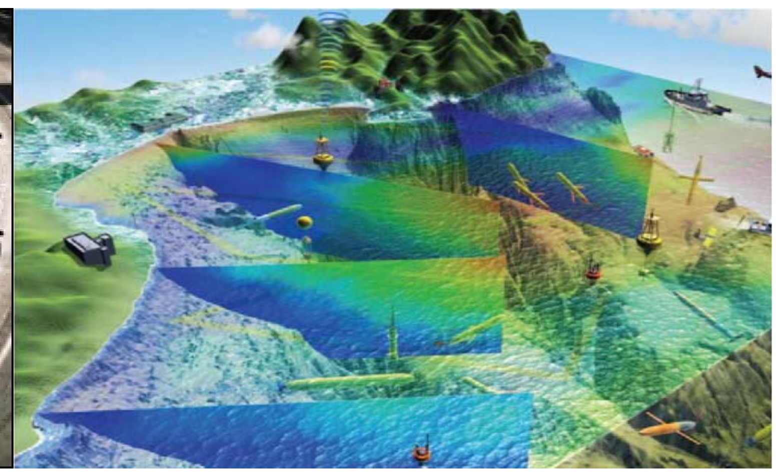
Collaborative Robotic Swarms



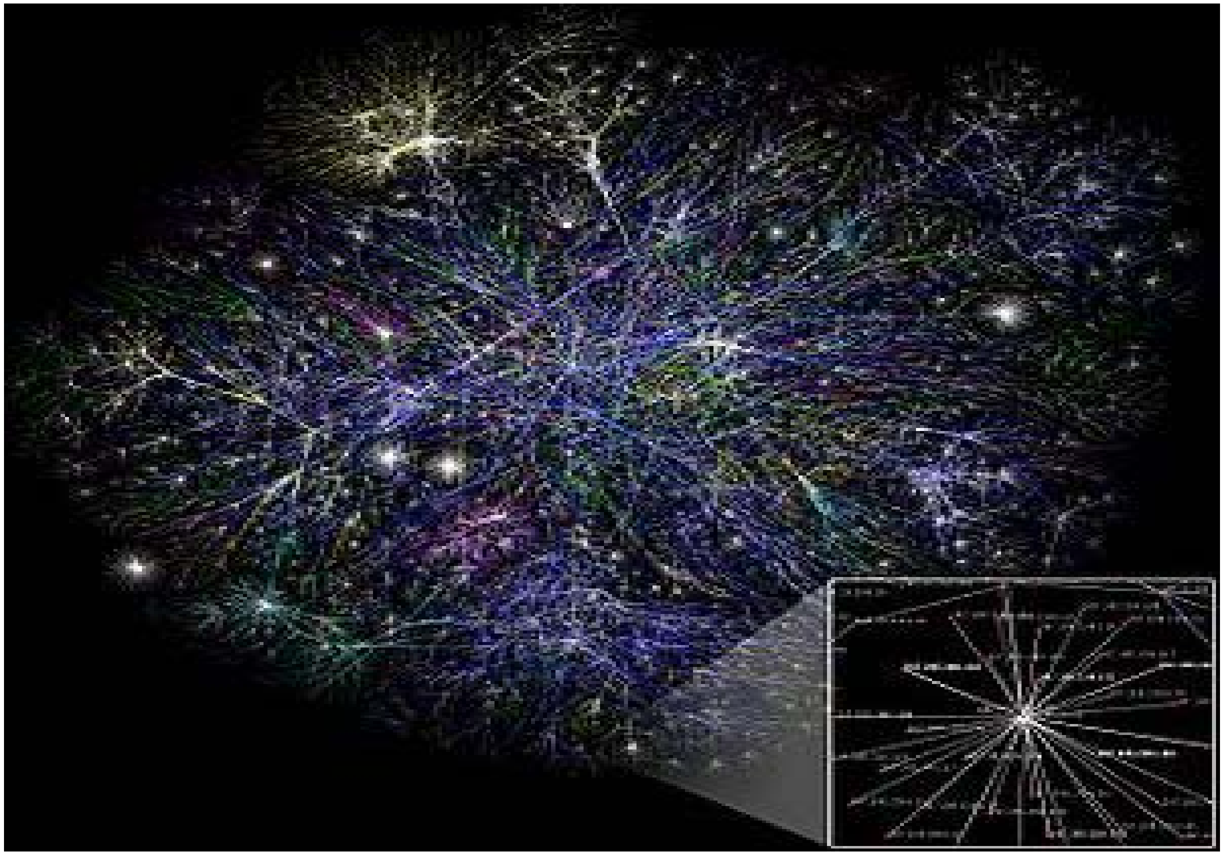
Autonomous Swarms – Networked Control



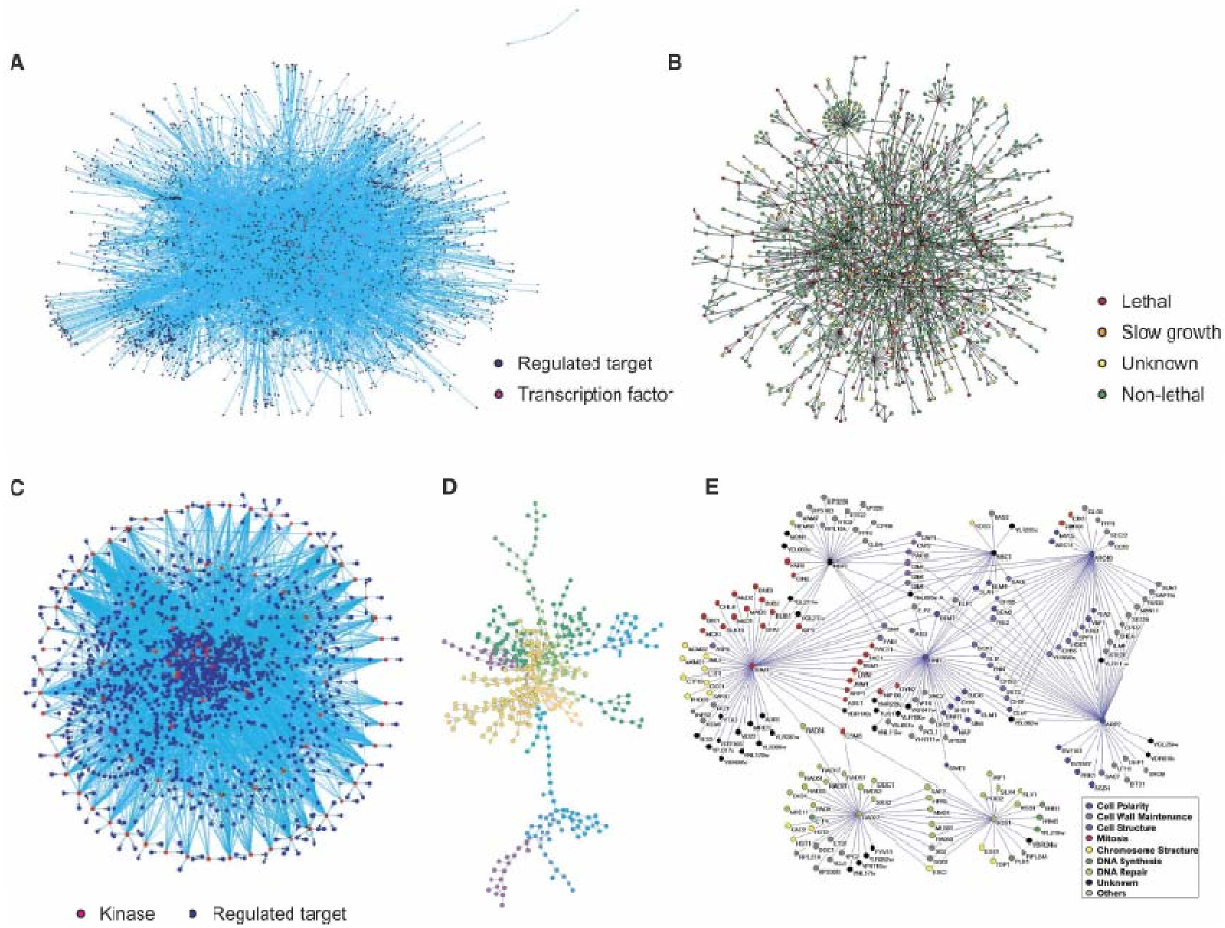




The Internet



Biological Networks

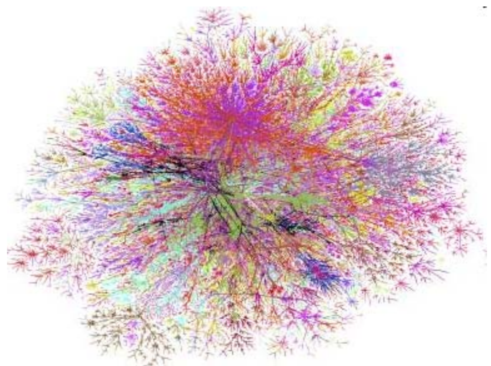


Examples of biological networks: [A] Yeast transcription factor-binding network; [B] Yeast protein -protein interaction network; [C] Yeast phosphorylation network ; [D] *E. Coli* metabolic network ; [E] Yeast genetic network ; Nodes colored according to their YPD cellular roles [Zhu et al, 2007]

The Institute for Systems Research

Networks and Networked Systems

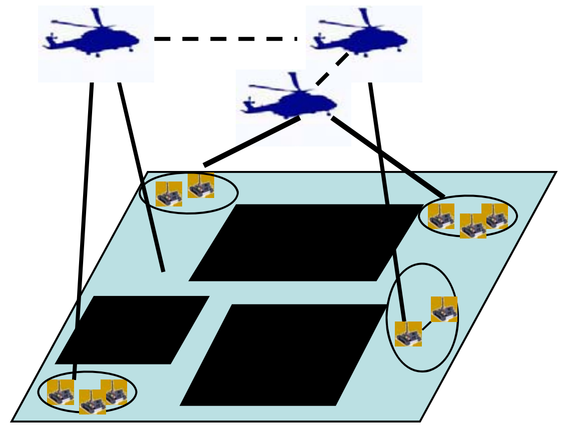
UNIVERSITY OF MARYLAND 1856



Internet backbone

(Lumeta Corp.)

Physical

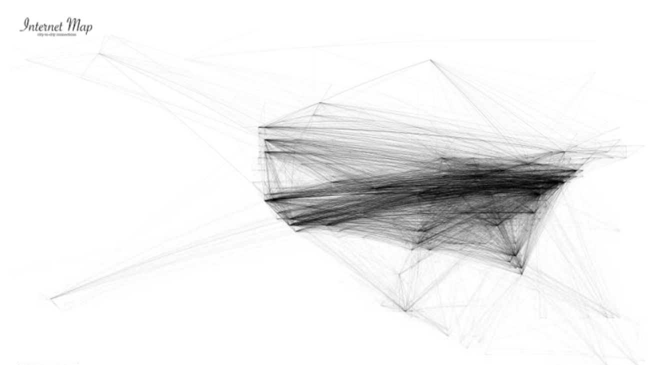


Vehicle, robot networks

Trust

(J Golbeck - Science, 2008)

Logical



Internet: North American cities

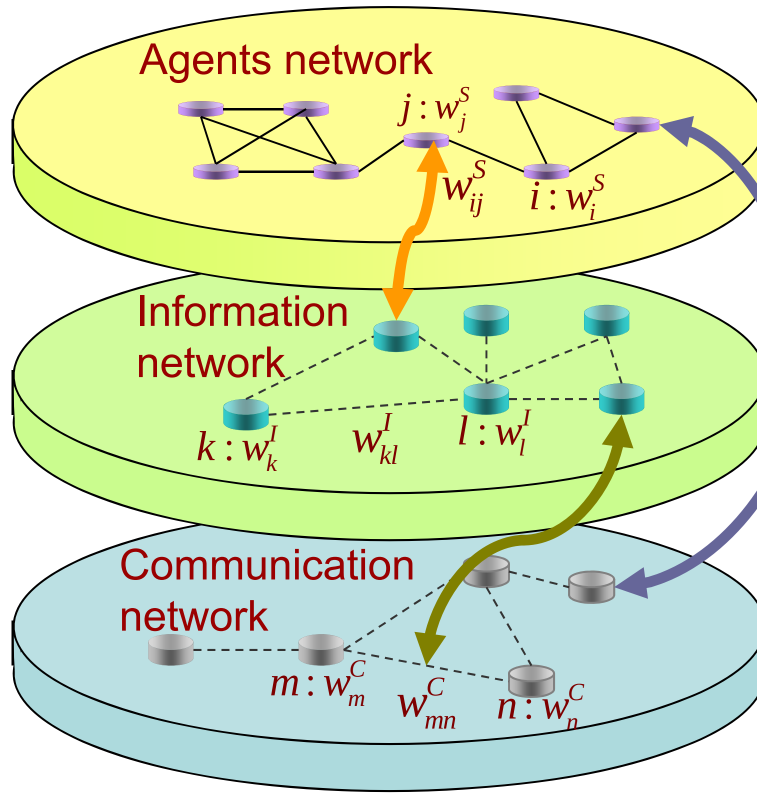
(Chris Harrison)

Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

Multiple Interacting Dynamic Hypergraphs

- Multiple Interacting Graphs
  - *Nodes*: agents, individuals, groups, organizations
  - Directed graphs
  - *Links*: ties, relationships
  - Weights on links : value (strength, significance) of tie
  - Weights on nodes : importance of node (agent)
- **Value directed graphs with weighted nodes**
- **Real-life problems: Dynamic, time varying graphs, relations, weights, policies**



Networked System architecture & operation

Network Complexity: Four Fundamental Challenges

- **Multiple interacting dynamic hypergraphs involved**
  - ***Collaboration hypergraph***: who collaborates with whom / when
  - ***Communication hypergraph***: who communicates with whom / when
- **Effects of connectivity topologies**:
  Find graph topologies with favorable tradeoff between performance (**benefit**) *vs* **cost** of collaborative behaviors
  - **Small word graphs** achieve such **tradeoff**
- Components, Interfaces, **Compositional Synthesis**
  - Network protocols – **component based networing**
  - Compositional Universal Security
- Need for **different probability models** – the classical Kolmogorov model is **not correct**
  - Probability models over logics and timed structures
  - **Logic of projections in Hilbert spaces** – not the Boolean of subsets

Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

What is a Network …?

- **In several fields or contexts:**
  - **social**
  - **economic**
  - **communication**
  - **sensor**
  - **biological**
  - **physics and materials**

A Network is …

- A collection of nodes, agents, …

  that **collaborate** to accomplish actions, gains, …

  that cannot be accomplished with out such collaboration

- Most significant concept for **dynamic autonomic networks**

The Fundamental Trade-off

- The nodes **gain** from collaborating
- But collaboration has **costs** (e.g. **communications)**
- **Trade-off: gain from collaboration *vs* cost of collaboration**

  Vector metrics involved typically

  Constrained Coalitional Games

- **Example 1**: **Network Formation** -- Effects on Topology
- **Example 2**: **Collaborative robotics, communications**
- **Example 3**: **Web-based social networks and services**
- **Example 4**: **Groups of cancer tumor or virus cells**

• • •

Example: Autonomic Networks

- Autonomic: self-organized, distributed, unattended
  - Sensor networks
  - Mobile ad hoc networks
  - Ubiquitous computing
- Autonomic networks depend on **collaboration** between their nodes for all their functions
  - The nodes gain from collaboration: e.g. multihop routing
  - Collaboration introduces cost : e.g. energy consumption for packet forwarding

Example: Social Webs

- In August 2007, there were totally 330,000,000 unique visits to social web sites. (Source: Nielsen Online)
  - 9 sites with over 10,000,000 unique visits
  - MySpace, Facebook, Windows Live Spaces, Flickr, Classmates Online, Orkut, Yahoo! Groups, MSN Groups

- Main types of social networking services
  - directories of some categories: e.g. former classmates
  - means to connect with friends: usually with self-description pages
  - **recommender systems** linked to **trust/reputation**

Gain

- Users gain by joining a coalition
  - **Wireless networks**
    - The benefit of nodes in wireless networks can be the rate of data flow they receive, which is a function of the received power

$$B_{ij} = f(P_j l(d_{ij}))$$

P_j is the power to generate the transmission and $l(d_{ij}) < 1$ is the loss factor

e.g: $B_{ij} = \log(1 + (P_j l(d_{ij}) / N_0))$

  - **Social connection** model (Jackson & Wolinsky 1996)

$$B_{ij} = \sum_{j \in g} V \delta^{r_{ij} - 1} \quad \text{or} \quad w_i(G)$$

    - r_{ij} is # of hops in the shortest path between i and j
    - $0 \leq \delta \leq 1$ *is the connection gain depreciation rate*

Cost

- Activating links is costly. $c_i(G) = \sum_{j \in N_i^t} C_{ij}$
  - **Wireless networks**
    - Energy consumption for sending data: $C_{ij} = RSd_{ij}^{\alpha}$

      RS depends on transmitter/receiver antenna gains and system loss not related to propagation

      α : path loss exponent
    - Data loss during transmission

      v_i is the environment noise and I_{ij} is the interference

      $$C_{ij} = h(v_i, I_{ij}) > 0$$
  - **Social connection model**
    - The more a node is **trusted**, the lower the cost to establish link

      e.g.suppose that the trust i has on j is s_{ij} (between 0 and 1), we can define the cost as the inverse of the trust values

      $$C_{ij} = 1 / s_{ij}$$

Coalitional Games

- **Payoff** of node $\boldsymbol{i}$ from the network G is defined as

$$v_i(G) = \text{gain} - \text{cost} = w_i(G) - c_i(G)$$

- **Iterated process**
  - Node pair $\boldsymbol{ij}$ is selected with probability $\boldsymbol{p_{ij}}$
  - If link $\boldsymbol{ij}$ is already in the network, the decision is whether to sever it, and otherwise the decision is whether to activate the link
  - The nodes act **myopically**, activating the link if it makes each at least as well off and one strictly better off, and deleting the link if it makes either player better off
  - **End**: if after some time, no additional links are formed or severed
  - **With random mutations** , the game converges to a unique Pareto equilibrium (underlying Markov chain states)

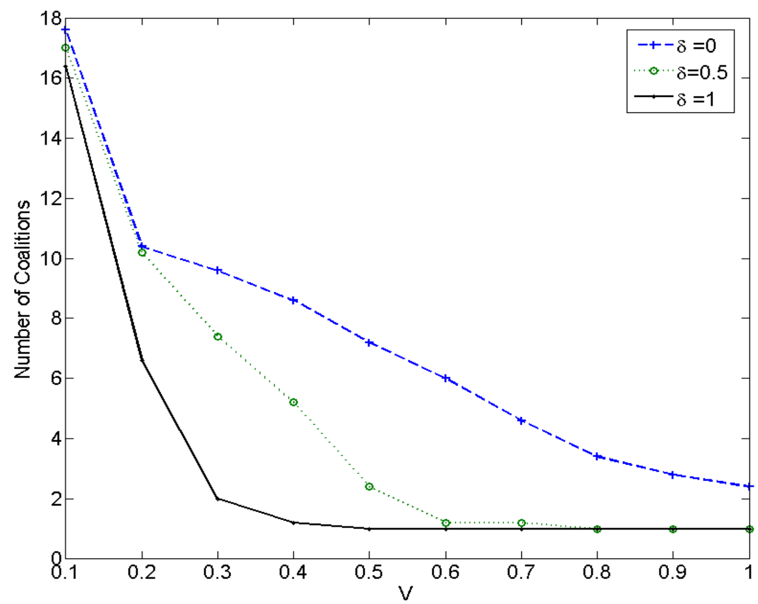
Stochastic Stability

- Dynamic process is now a finite state, aperiodic, irreducible Markov chain (graph process)-- steady-state distribution, $\Pi(g, \varepsilon)$.
- A network *g* is **stochastically stable** if $\Pi(g, \varepsilon)$ is bounded below as the error rate, ε, tends to zero;

  $\Pi(g, \varepsilon) \rightarrow a > 0$, as $\varepsilon \rightarrow 0$.
  - Stochastically stable networks must be pairwise stable networks or networks of closed cycles
  - Stochastic stability identifies the most “robust” or easy to reach networks in a particular sense (the most mutations needed to get “unstuck”).
  - The above example converges to a Pareto efficient pairwise stable network by considering all the possible dynamic paths between the left and right networks.

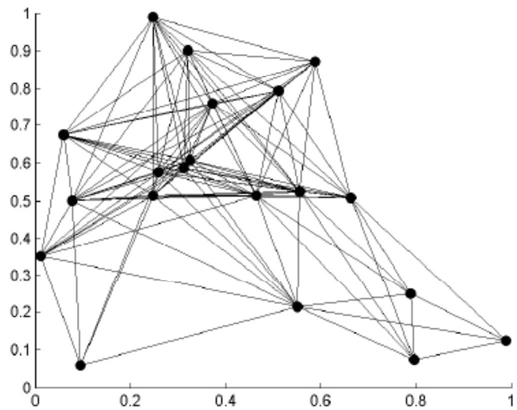
Coalition Formation at the Stable State

- The cost depends on the physical locations of nodes
  - Random network where nodes are placed according to a uniform Poisson point process on the [0,1] x [0,1] square.
- Theorem: The coalition formation at the stable state for n→∞
  - Given $\delta = 0,\quad V = P\left(\frac{\ln n}{n}\right)^{\frac{\alpha}{2}}$ is a sharp threshold for establishing the grand coalition (number of coalitions = 1).
  - For $0 < \delta \leq 1$, the threshold is less than $P\left(\frac{\ln n}{n}\right)^{\frac{\alpha}{2}}$.

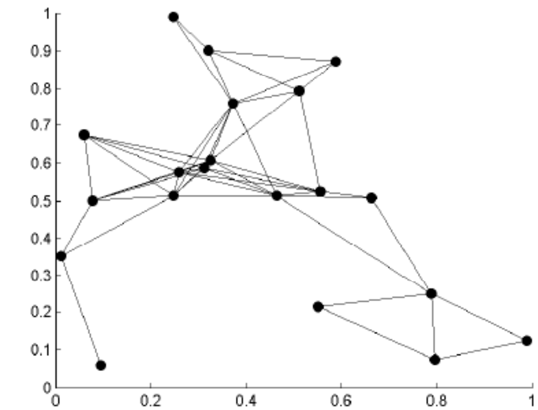


n = 20

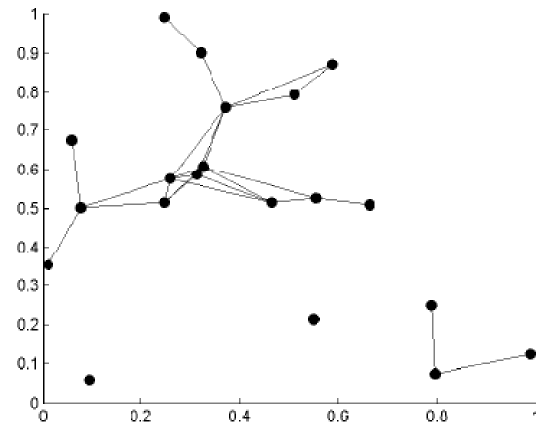
Topologies Formed



(a) P = 0.5 (low cost); complete graph



(b) P = 2 (middle cost); small world topology



(c) P = 4 (high cost); partitioned network

Time-dependent Game

- The game is time-dependent
  - The payoff players receive **varies over time**.
  - The dynamics of the game can be separated in rounds of successive coalition expansions (or contractions).

- The dynamic coalition formation process is described as an iterated game
  - x_i^t : the action *i* chooses at time *t*.
  - $v_i\left(x^t\right)$: the payoff of user *i* at time *t.*
  - $q^t(x)$: players' probability of playing action *x* at time *t.*
  - C_i^t : the set of users that form the coalition user *i* belongs to at time *t*.
  - user *i* and user *j* decide to activate link *ij* at time *t:*

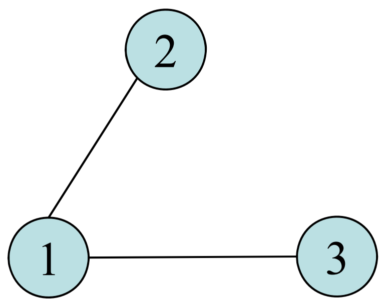
$$C_i^t = C_j^t = C_i^{t-1} \bigcup C_j^{t-1}$$

Value Function

- Value function for coalition *C* (component-wise additive value function)

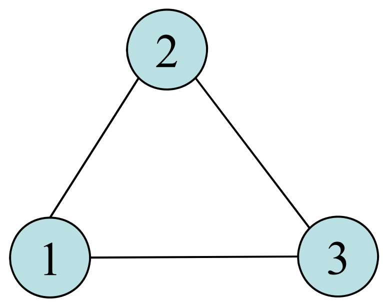
$$v(C) = \sum_{i \in C} v_i(g)$$

- Value function depends on topology



Same coalition $C=\{1,2,3\}$
with different topology

$v(\{12,13\}) \neq v(\{12,23,13\})$



Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

Networks and Trust

- **Trust and reputation critical for collaboration**
- Characteristics of trust relations:
  - *Integrative* (Parsons1937) – main source of social order
  - *Reduction of complexity* – without it bureaucracy and transaction complexity increases (Luhmann 1988)
  - *Trust as a lubricant for cooperation* (Arrow 1974) – rational choice theory
- **Social Webs, Economic Webs**
  - MySpace, Facebook, Windows Live Spaces, Flickr, Classmates Online, Orkut, Yahoo! Groups, MSN Groups
  - e-commerce, e-XYZ, services and service composition
  - **Reputation** and **recommender** systems

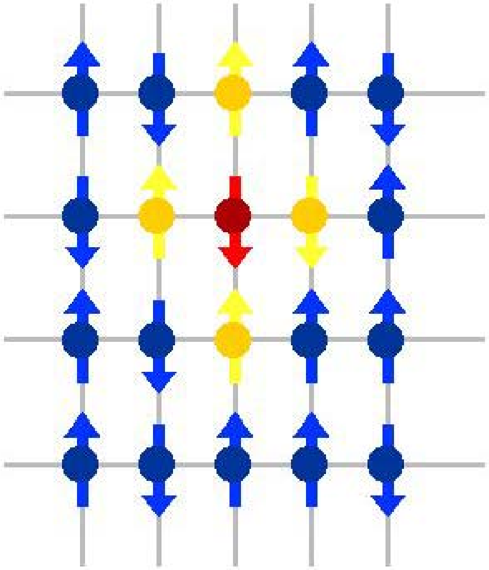
Ising and Spin Glass Models

- Statistical Physics models for magnetization

  Orientation of each particle's spin depends on its **neighbors**

  Ising Model: behavior of simple magnets

  Spin Glass Model: complex materials



- Interpretation:

  $\boldsymbol{s} = \{s_1, s_2, \ldots, s_n\}$ is a configuration of n particle spins -- $s_j = 1$ or -1 (up or down)

  Energy for configuration $\boldsymbol{s}$

$$H(\mathbf{s}) = -\frac{1}{T}\sum_{\substack{i \in V \\ j \in N_i}} J_{ij} s_i s_j - \frac{mH}{T}\sum_i s_i$$

  – **Ising Model**: $J_{ij} = J$ for all i, j

  – **Spin Glass Model**: J_{ij} depend on i,j and can be **random**

Ising/SG Models and Games

- Ising/SG models can be interpreted as dynamic (repeated) games:
  - The value of s_i represents whether node i is willing to cooperate or not
  - each particle selects spin to **maximize** its own *payoff*

$$\pi_i = \left(\sum_{j \in N_i} J_{ij} s_i s_j\right) / T$$

  - Ising model ($J_{ij} = J > 0$) : align its spin with the majority of neighbors spin
    - High T, conservative agents, not willing to change, small payoffs
    - Low T, aggressive agents, larger payoffs
  - Collection of local decisions reduces the total energy of the interacting particles

- Inspires an approach where **trust** is an **incentive for cooperation**
  - J_{ij} can be interpreted as the **worth of player *j* to player *i***
  - decide to cooperate or not based on benefit from cooperation **and** trust values of neighbors

Trust as Mechanism to Induce Collaboration

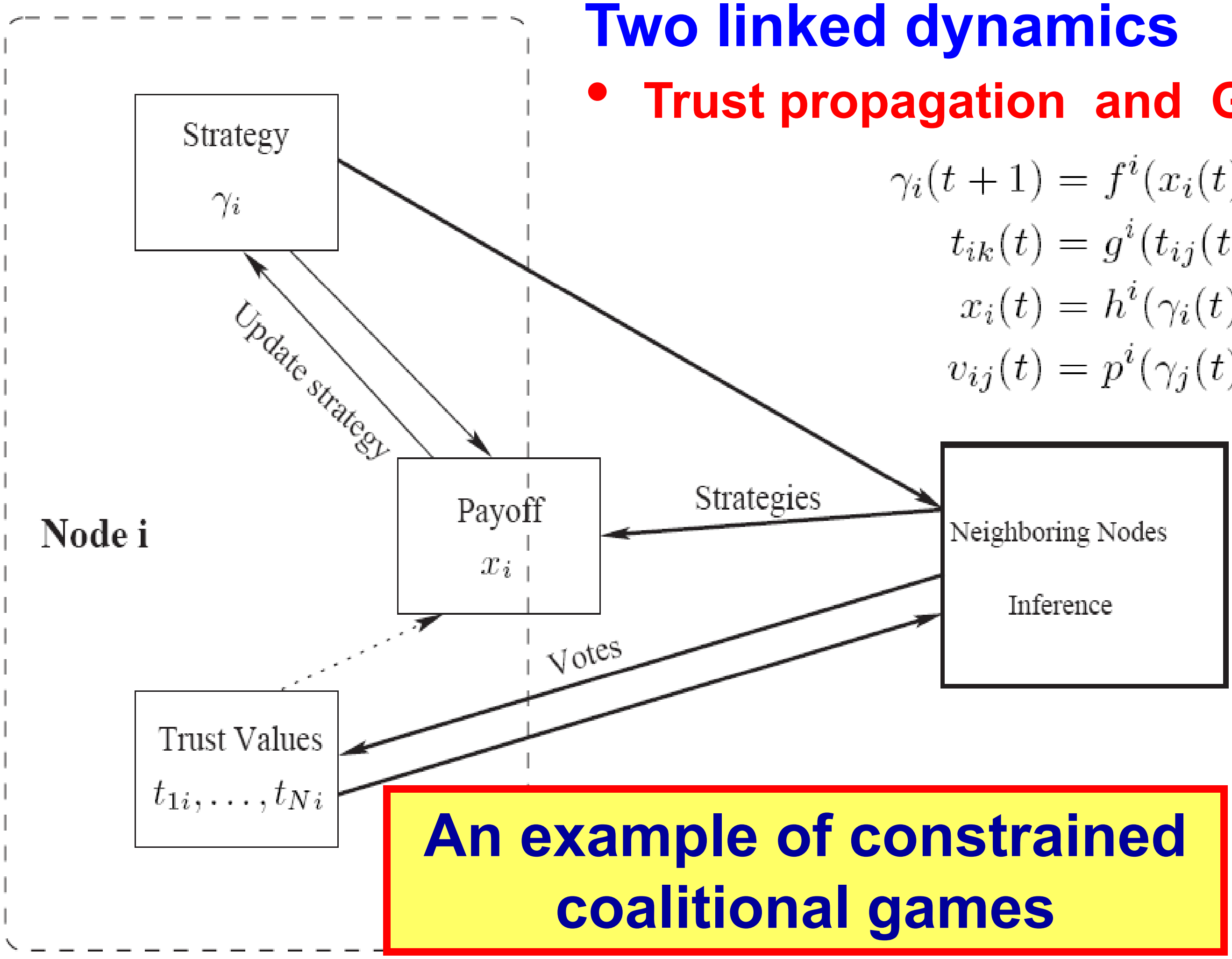
- Trust **is an incentive** for collaboration
  - Nodes who refrain from cooperation get lower trust values
  - Eventually penalized because **other nodes tend to only cooperate with highly trusted ones**.
- For node i **loss for not cooperating** with node j is a nondecreasing function of J_{ji}, $f(J_{ji})$,
- New characteristic function is

$$\nu(\mathcal{S}) = \sum_{i,j\in\mathcal{S}} J_{ij} - \sum_{i\in\mathcal{S},j\notin\mathcal{S}} f(J_{ij})$$

- **Theorem** : if $\forall i, j,\ J_{ij} + f(J_{ji}) \geq 0$, the core is nonempty and $x_i = \sum_{j\in N_i} J_{ij}$ is a feasible payoff allocation in the core.

By introducing a trust mechanism, all nodes are induced to collaborate without any negotiation

Dynamic Coalition Formation



Two linked dynamics

- **Trust propagation and Game evolution**

$$\gamma_i(t+1) = f^i(x_i(t), \gamma_i(t), \gamma_j(t), t_{ij}(t))$$
$$t_{ik}(t) = g^i(t_{ij}(t), v_{jk}(t)) \quad \forall k \in N$$
$$x_i(t) = h^i(\gamma_i(t), \gamma_j(t))$$
$$v_{ij}(t) = p^i(\gamma_j(t), t_{ji}(t))$$

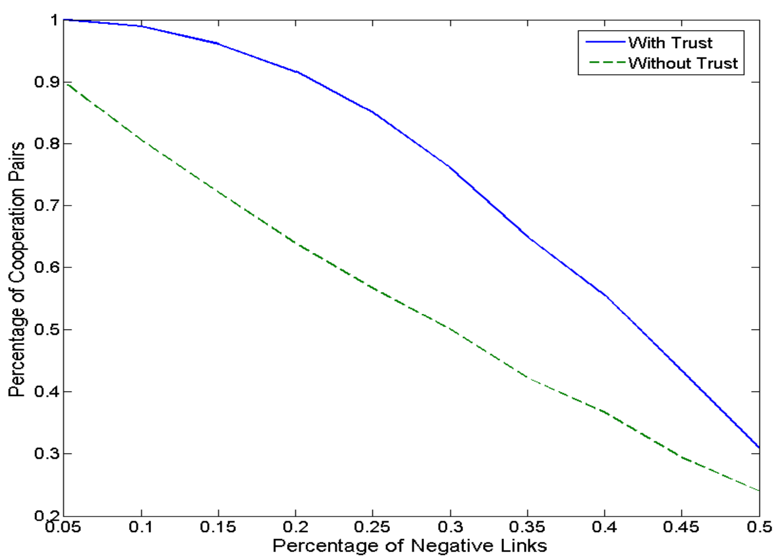
An example of constrained coalitional games

Stability of dynamic coalition

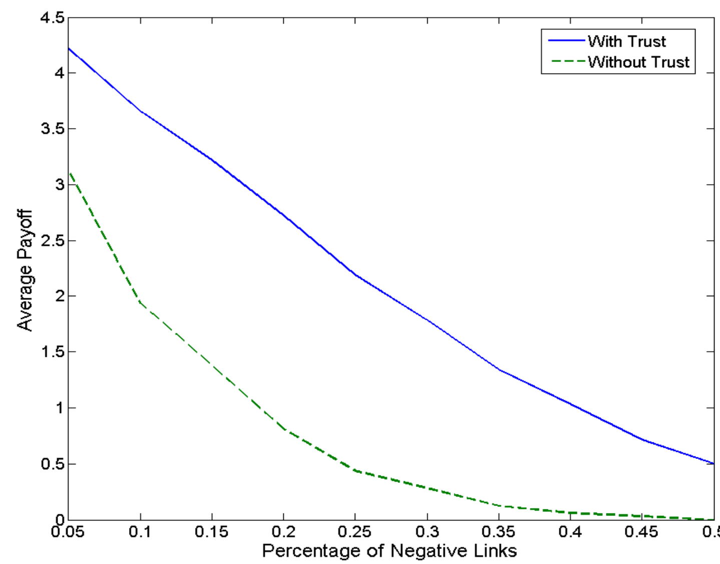
Nash equilibrium

Results of Game Evolution

- **Theorem**: $\forall i \in N_i$ and $x_i = \sum_{j \in N_i} J_{ij}$, there exists τ_0, such that for a reestablishing period $\tau > \tau_0$
  - terated game converges to Nash equilibrium;
  - In the Nash equilibrium, all nodes cooperate with all their neighbors.
- Compare games with (without) trust mechanism, strategy update:



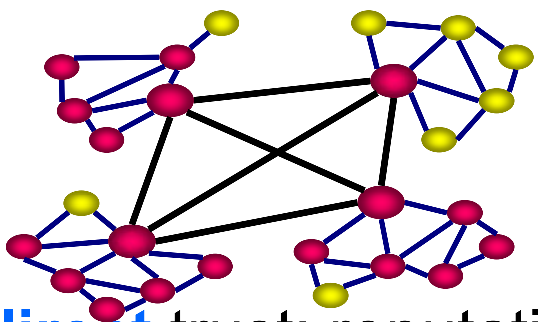
Percentage of cooperating pairs *vs* negative links



Average payoffs *vs* negative links

Next Generation Trust Analytics

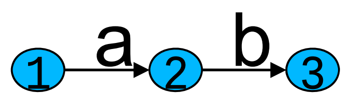
- Trust evaluation, trust and mistrust dynamics
  - Spin glasses (from **statistical physics**), phase transitions



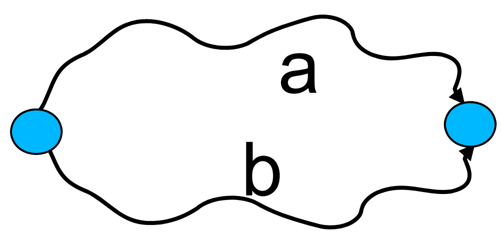
$$s_i(k+1) = f\left(\hat{J}_{ji}, s_j(k) \mid j \in N_i\right)$$

- **Indirect** trust; reputations, profiles; Trust computation via 'linear' **iterations in ordered semirings**

$$a \otimes b \leq a, b$$



$$a \oplus b \geq a, b$$



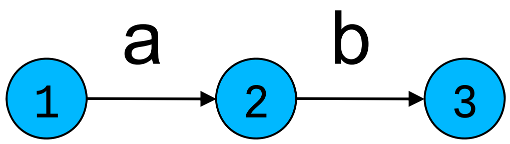
2007 IEEE Leonard Abraham prize
New Book **"Path Problems in Networks" 2010**

- **Direct trust: Iterated pairwise games on graphs** with players of many types

Generalized Networks and Semirings

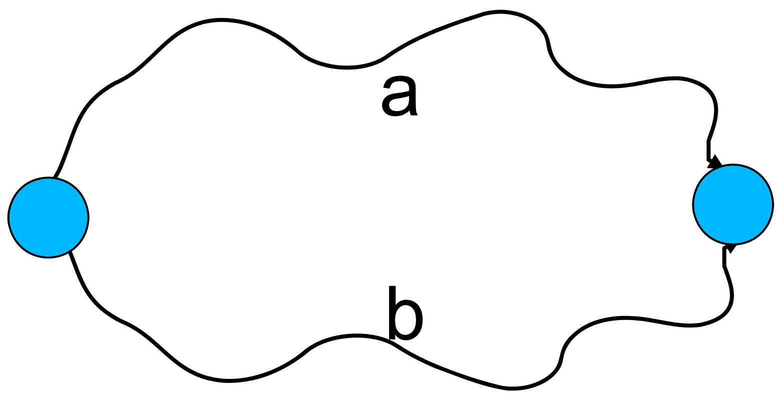
- Combined **along-a-path weight should not increase** :

$$a \otimes b \leq a, b$$



- Combined **across-paths weight should not decrease** :

$$a \oplus b \geq a, b$$



Distributed Weight Dynamics

- Path interpretation

$$t_{i \to j} = \bigoplus_{\text{path } p: i \to j} t^{p}_{i \to j}$$

- **Linear system interpretation**

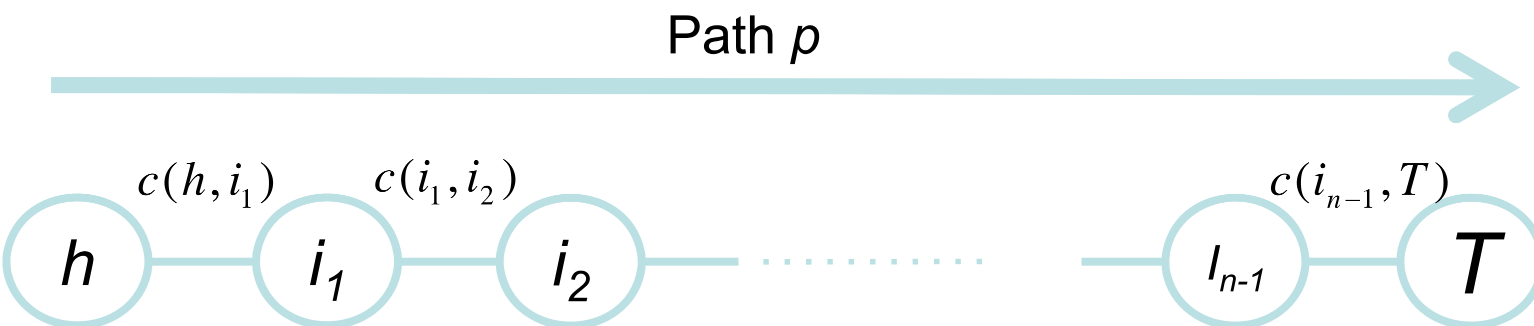
$$t_{i \to j} = \bigoplus_{User\ k} t_{i \to k} \oplus w_{k \to j}$$

$$\vec{t}_n = W \otimes \vec{t}_{n-1} \oplus \vec{b}$$

- Treat as a **linear system**
  – We are looking for its **steady state**.
- Have resolved also **shortest path sensitivity** in semiring framework

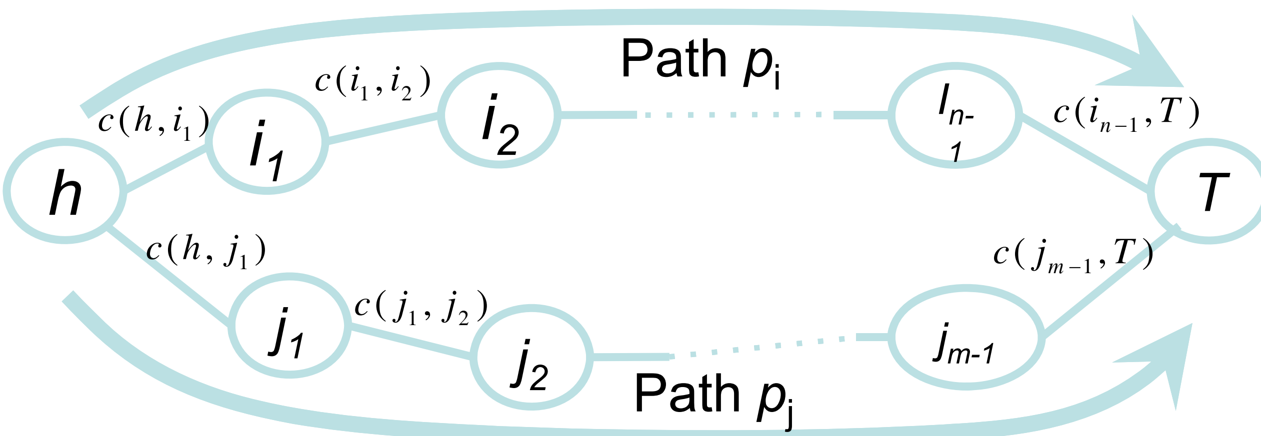
Graphs and Semirings

Consider undirected graph $G(V,E)$ with edge weights $c(i,j) \in S$.



Weight of the path p by edge composition is

$$w(p) = c(h,i_1) \otimes c(i_1,i_2) \otimes \ldots \otimes c(i_{n-1},T)$$



Composing path weights:

$$w(p_i) \oplus w(p_j)$$

Ordered Semirings

- Functions with semiring structure lend themselves to distributed computation/evaluation
- **Ordered semirings** $\oplus$ is the supremum or infimum operator and $(S, \otimes, \preceq)$ is an ordered semigroup
- Associate with every edge (*i*, *j*) of the dynamic graph a semiring element $c(i,j)[t] \in S$
- **General semiring optimal path problem** on a dynamic graph corresponds to computing

$$p^*(S,T)[t] = \arg\oplus_{p \in P_{S,T}[t]} \otimes_{(i,j)\in p} c(i,j)[t]$$

Multi-Criteria Shortest Path (several metrics)

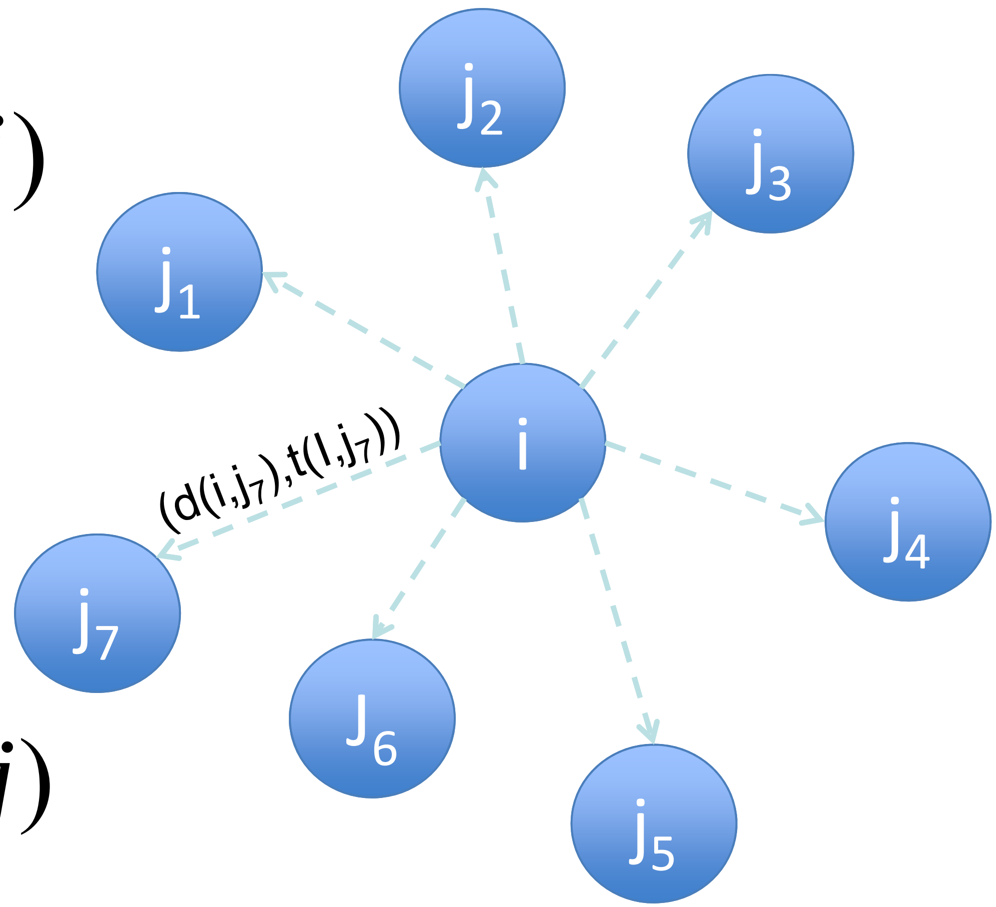
- **D** of a path "p"

$$d(p) = \sum_{(i,j)\in p} d(i,j)$$

- **T** of a path "p" – bottleneck

$$t(p) = \min_{(i,j)\in p} t(i,j)$$

Bi-metric Network



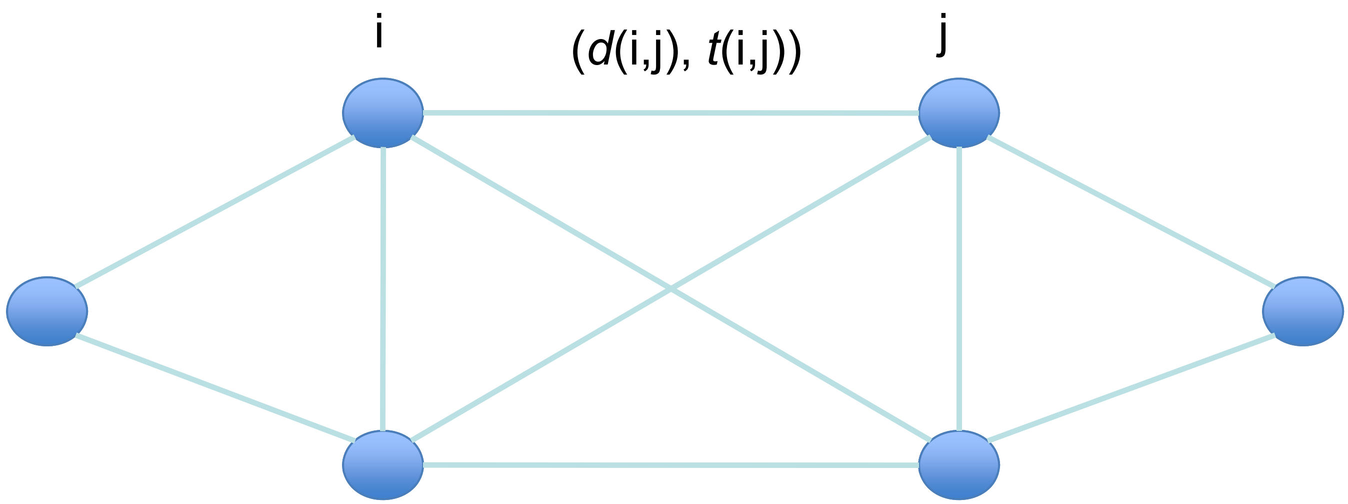
MCOP

- The two metrics are not trivially comparable.

$$MCOP: \quad (\mathcal{P}_{SD}, f, X) / \theta / (\mathcal{R}^{Q}, \preceq)$$

Notation	Definition	Name
$x \leq y$	$x_i \leq y_i \quad i = 1,2,..,Q$	Weak component-wise order
$x < y$	$x_i \leq y_i \quad i = 1,2,..,Q$ and $x \neq y$	Component-wise order
$x \ll y$	$x_i < y_i \quad i = 1,2,..,Q$	Strict component-wise order
$x \leq_{lex} y$	$x_k < y_k \quad or \quad x = y \quad k = \min\{i : x_i \neq y_i\}$	Lexicographic component-wise order
$x \leq_{MO} y$	$\max_i x_i \leq \max_i y_i$	Max order

Path problems on Graphs – D and T Semirings



$$\min_{p \in \mathcal{P}_{SD}} d(p) = \min_{p \in \mathcal{P}_{SD}} \sum_{(i,j) \in p} d(i,j)$$

$$\max_{p \in \mathcal{P}_{SD}} t(p) = \max_{p \in \mathcal{P}_{SD}} \left(\min_{(i,j) \in p} t(i,j) \right) = -\min_{p \in \mathcal{P}_{SD}} \left(\max_{(i,j) \in p} (-t(i,j)) \right)$$

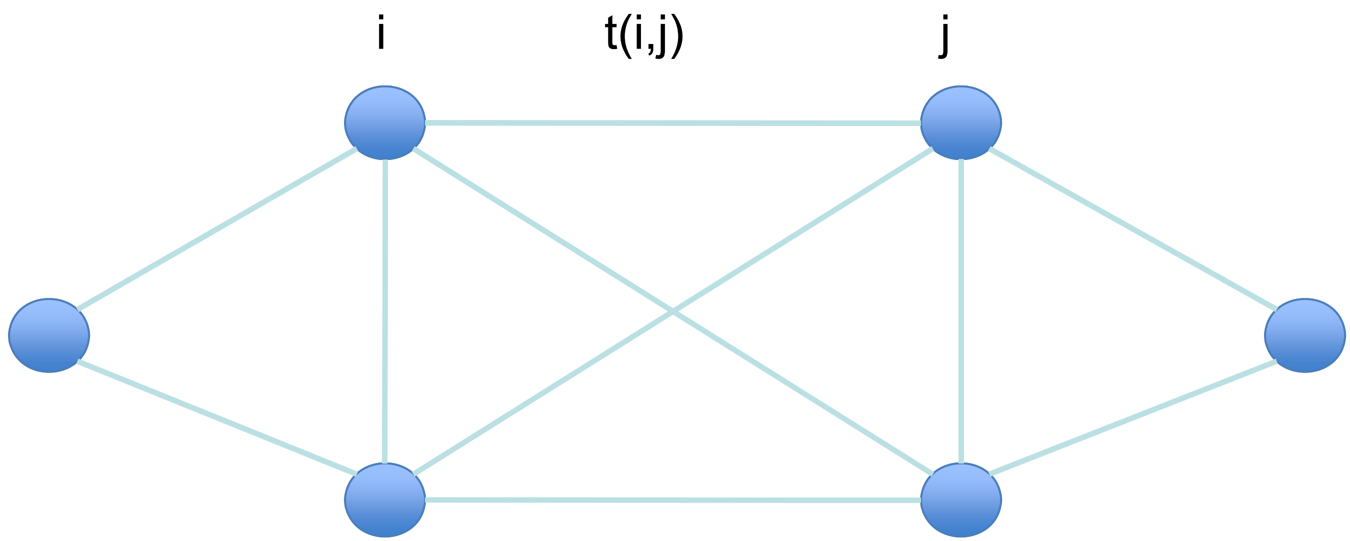
D Semiring: $(\mathcal{R}_+ \cup \{0\},\ \min,\ +)$

T Semiring: $(-\mathcal{R}_+ \cup \{0\},\ \min,\ \max)$

$$f : \mathcal{P}_{SD} \to R^2$$
$$f(p) = (d(p), -t(p)),\ \forall p \in \mathcal{P}_{SD}$$

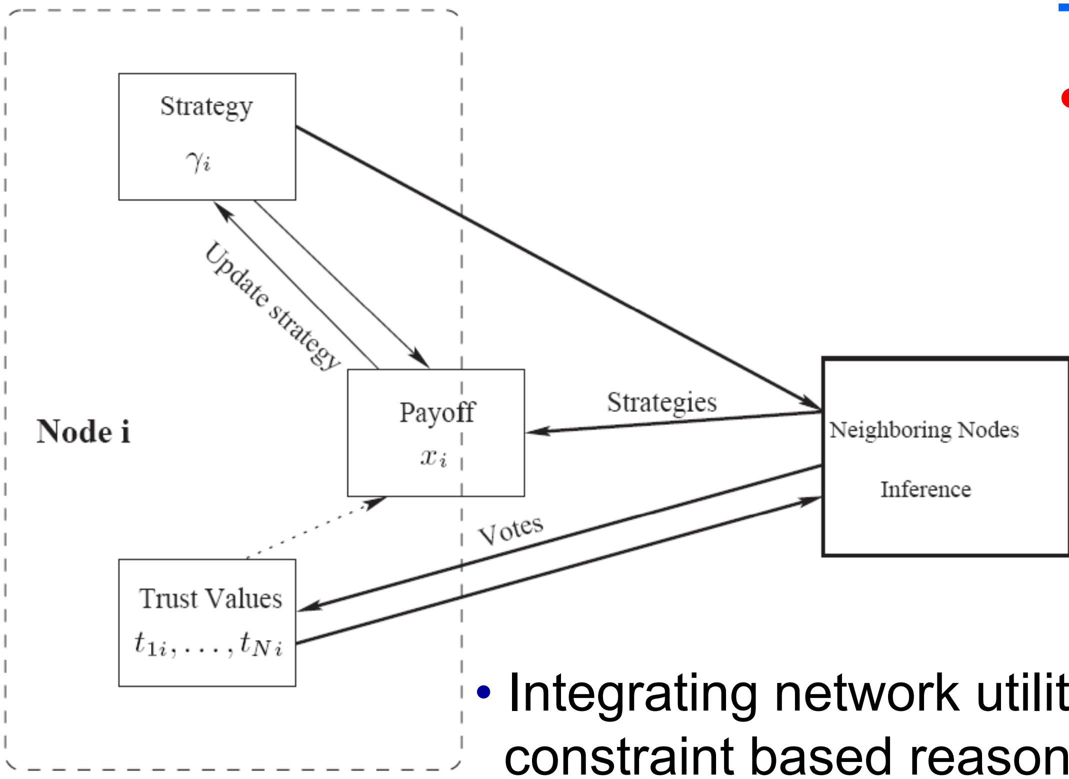
Notions of Optimality: Pareto, Lexicographic, Max-Ordering, Approx. Semirings

Pareto optimal paths – Edge exclusion algorithm



- Edge exclusion – From $G(V,E)$, remove all the edges whose $t(i,j) > \varepsilon$ to obtain a graph $G'(\varepsilon)$
- $G'(\varepsilon)$ contains paths which have all $t(i,j) \leq \varepsilon$
- We can also show that G' has all paths in G which have $t(i,j) \leq \varepsilon$ and only those

Constrained Coalitional Games: Trust and Collaboration



Two linked dynamics

- Trust / Reputation propagation and Game evolution

- Integrating network utility maximization (NUM) with constraint based reasoning and coalitional games

- Beyond linear algebra and weights, semirings of constraints, constraint programming, soft constraints semirings, policies, agents
- Learning on graphs and network dynamic games: behavior, adversaries
- Adversarial models, attacks, constrained shortest paths, …

Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

Component-Based Heterogeneous Network Synthesis

- **How to synthesize resilient, robust, adaptive networks?**

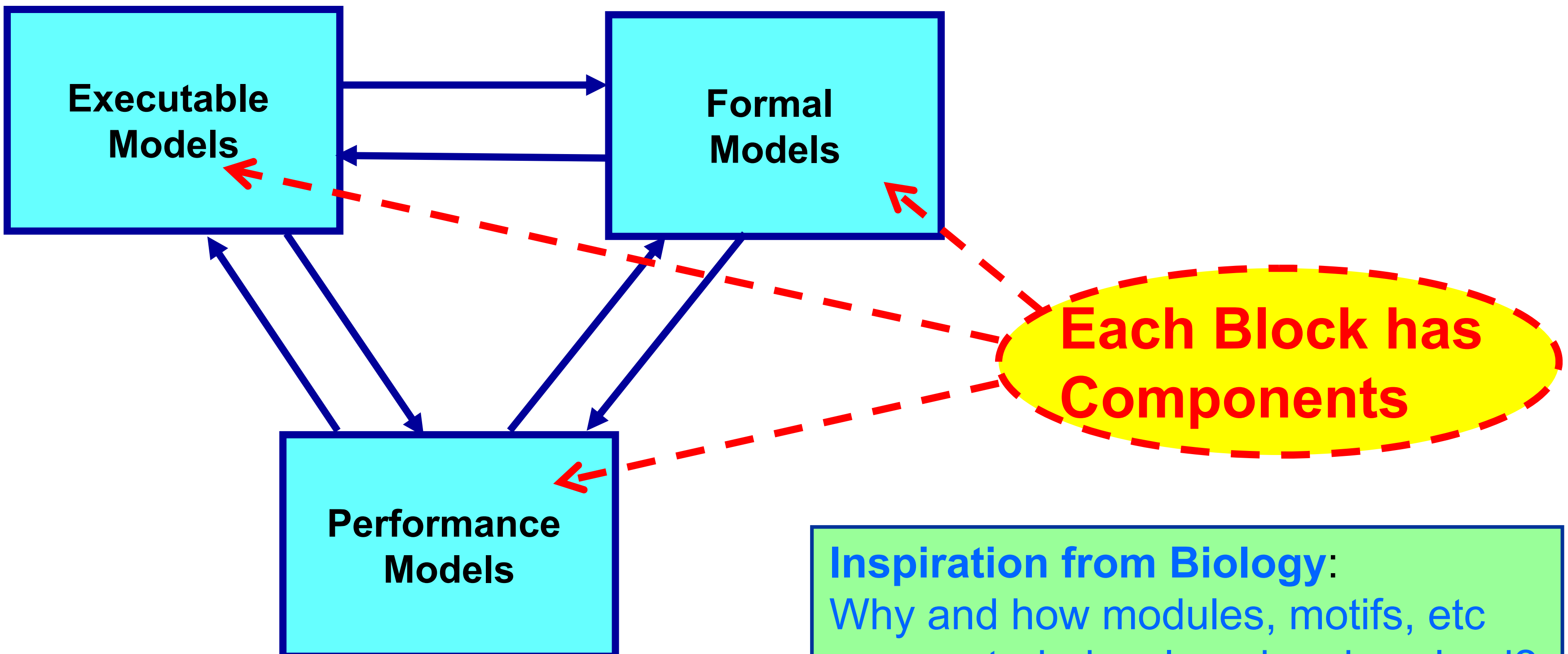
  → ***Component-Based Network Analysis & Synthesis (CBN)***
- **Components: modularity, cost reduction, re - usability, adaptability to goals, new technology insertion, validation and verification**
- **Interfaces: richer functionality– intelligent/cognitive networks**
- **Theory and Practice of Component-Based Networks**
  - **Heterogeneous components and compositionality**
  - **Performance of components and of their compositions**
  - **Back and forth from performance - optimization domain to correctness and timing analysis domain and have composition theory preserving component properties as you try to satisfy specs in both domains**
- **From communication to social, from cellular to transportation, from nano to macro networks**
- **Critical theory and methodology for Networked Embedded Systems, CyBer Physical Systems, Systems Biology**

Networks: Different Linked Views

Networks:

- as **distributed, asynchronous, feedback (many loops), hybrid automata (dynamical systems)**
- as **distributed asynchronous active databases** and **knowledge bases**
- as **distributed asynchronous computers**

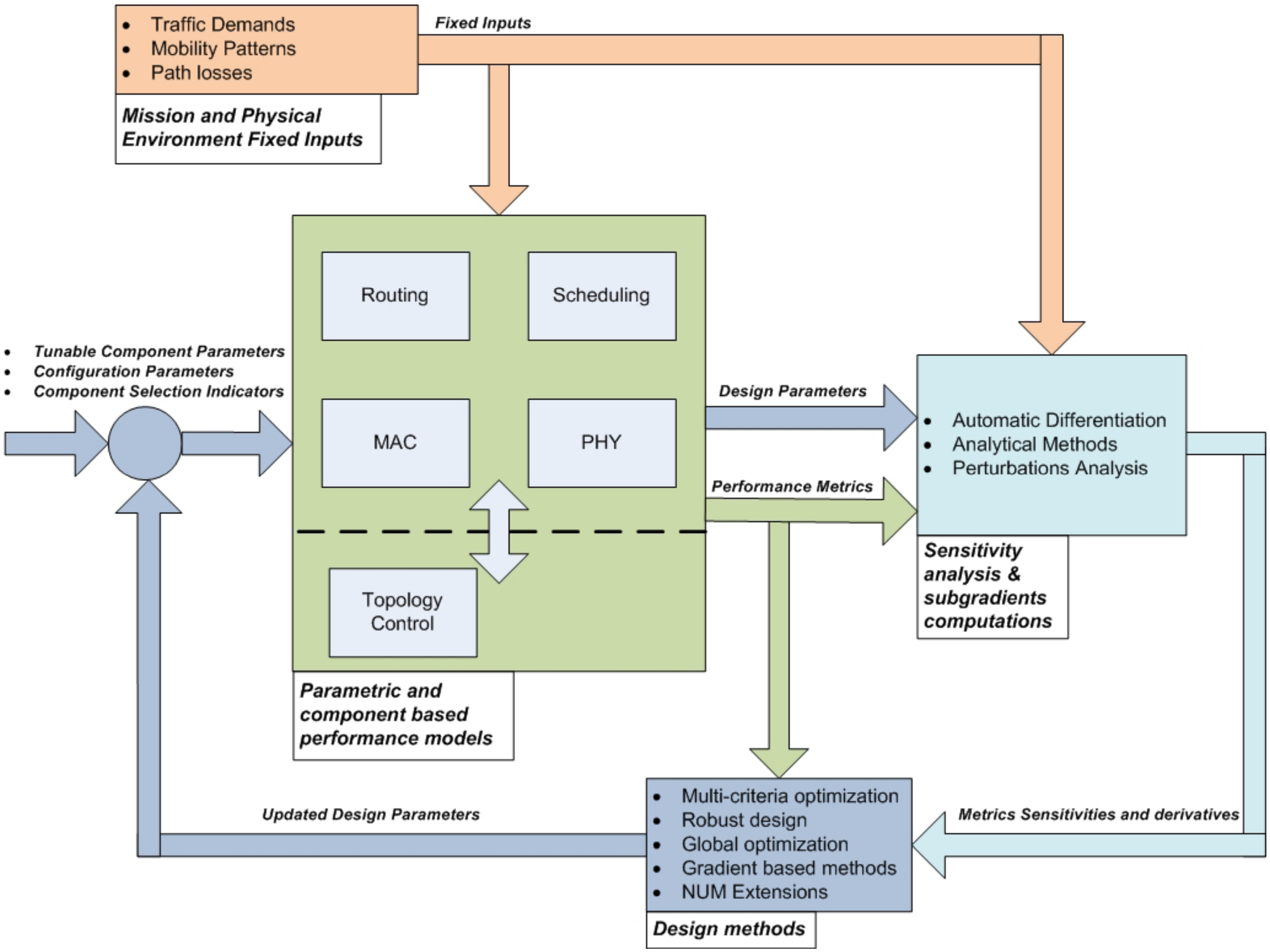
Component-Based Heterogeneous Networked Systems Synthesis



Inspiration from Biology: Why and how modules, motifs, etc are created, developed and evolved?

Grand challenge: Develop this framework for distributed, partially asynchronous systems, with heterogeneous components and time semantics

Component Based Routing (CBR) and Networking (CBN) for MANET



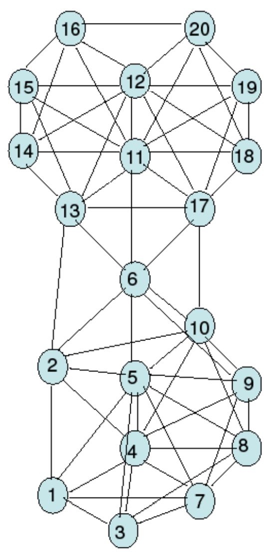
MAC and Routing Components

- **Objective**
  - Design MANET adaptable to missions with predictable performance
- **Approach**
  - **Break traditional layers to components!** Develop component-based models MANET that considers cross-layer dependency to improve the performance
  - Study the effect of each component on the overall MANET performance

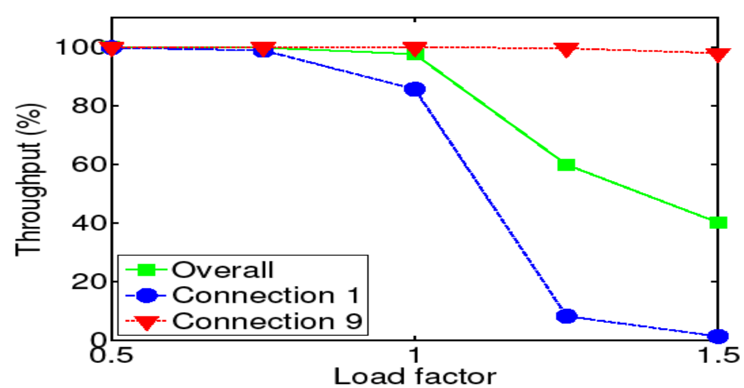
- **Routing Components** – routing protocols like OLSR [Baras08]
  - Neighbor Discovery Component (NDC)
  - Selector of Topology Information to Disseminate Component (STIDC)
  - Topology dissemination Component (TDC)
  - Route Selection Component (RSC)
- **MAC Components** – based on CSMA-CA MAC protocols like IEEE 802.11 [Baras08], and on schedules based MAC (USAP) [Baras09]
  - Scheduler
  - MAC

Results Todate

- Realistic MANET scenarios from DARPA CB-MANET benchmarks
- 20 node, 10 connections -- up to 50 moving nodes with disconnections (ground and UAVs)
- **Substantial improvements of performance through new NDC and STIDC** components (as compared to traditional OLSR) – being reported to MANET WG of IRTF



PHY Layer Connectivity



Throughputs for increasing load

Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

Distributed Algorithms in Networked Systems and Topologies

- Distributed algorithms are essential
  - Group of agents with certain abilities
  - Agents **communicate with neighbors,** share/process information
  - Agents **perform local** actions
  - **Emergence** of global behaviors
- **Effectiveness** of distributed algorithms
  - The **speed** of convergence
  - **Robustness** to agent/connection failures
  - Energy/ communication **efficiency**
- **Group topology affects** group performance
- **Design problem**:

  Find graph topologies with favorable tradeoff between performance improvement (**benefit**) *vs* **cost** of collaboration
- **Example: Small Word graphs** in consensus problems

Design of information flow

Consensus problems

$$x(k+1) = F(k)x(k)$$

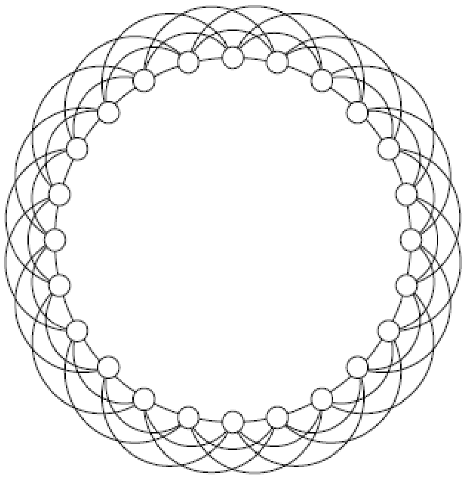
$$F(k) = (I + D(k))^{-1}(A(k) + I)$$

$$F(k) = I - hL(k)$$

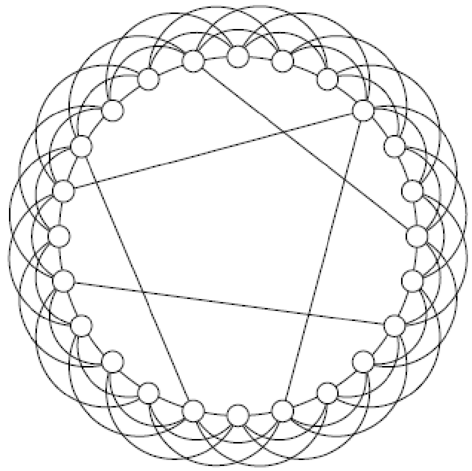
Symmetric communication

- **Fixed graphs**: Geometric convergence with rate equal to Second Largest Eigenvalue Modulus (SLEM)
- How does **graph topology affect** location of eigenvalues?
- How can we **design graph topologies** which result in good convergence speed?

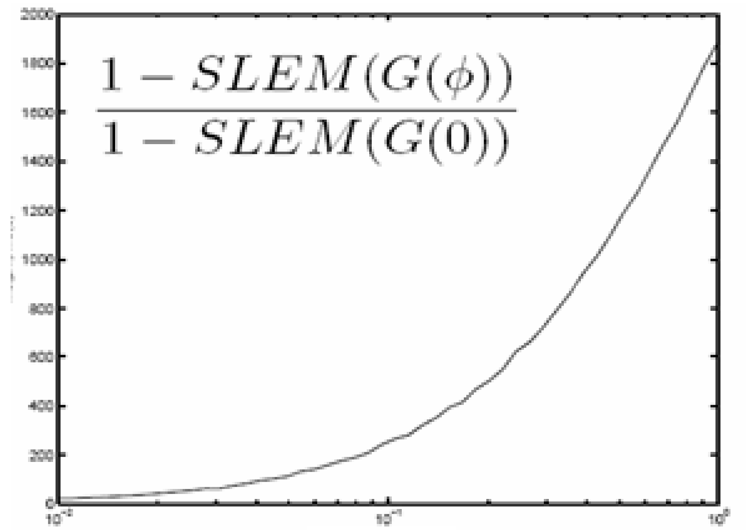
Small World Graphs



Simple Lattice
C(n,k)



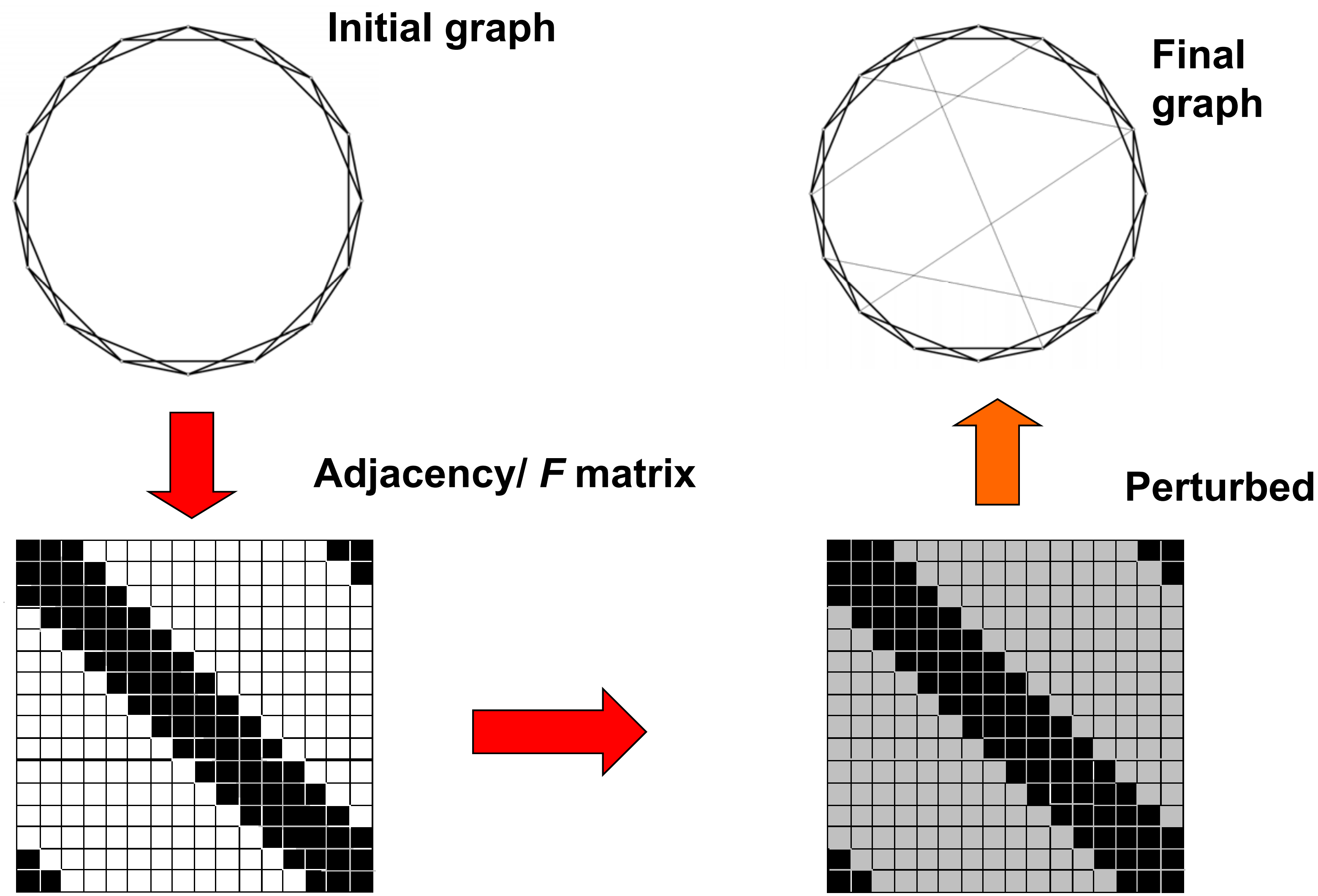
Small world: Slight variation adding $nk\Phi$



\phi
C(1000,5)

Adding a **small portion** of well-chosen links → **significant increase** in convergence rate

Mean Field Explanation and Perturbation Approach



Watts-Strogatz Small World networks

- Random graph approach
  (e.g. Durrett 2007, Tahbaz and Jadbabaie 2007)
- **Perturbation** approach (Higham 2003)
  - Start from lattice structure $G_0 = C(n,k) \longleftrightarrow F_0$
  - Perturb zero elements in the positive direction by $\varepsilon = \frac{K}{n^{\alpha}}$ for fixed $K > 0$ and $\alpha > 1$.
  - Perturb the formerly nonzero elements equally, such that the stochastic structure of the *F* matrix is preserved F_{ε}
  - Analyze the SLEM as a function of the perturbation as α varies

1- D case …

- Refer to the perturbations as ***ε*-shortcuts**
- In the limit of large *n* :
  - For $\alpha > 3$ the effect of *ε*-shortcuts on convergence rate is negligible
  - For $\alpha = 3$ the effect of ε-shortcuts on convergence rate starts (spectral gap gain perturbation of same order)
  - **For $\alpha = 2$ the shortcuts dominantly decrease SLEM**
  - For $\alpha = 1$ SLEM is very small
- *ε*-shortcuts are loosely analogous to the shortcuts in Small World networks
- ***α* = *3*** can be considered as the onset of small world effect with **small world effect happening** at ***α* = 2**

Distributed exploration of the graph structure

- **Self-organization for better performance and resiliency**
- Hierarchical scheme to design a network structure capable of running **distributed algorithms with high convergence speed**
- A two stage algorithm:

  1- Find the most effective choice of **local leaders**

  2- Provide nodes with information about their location **with respect to other nodes and leaders** and the choice of groups to form
- Divide N agents into K groups with M members each

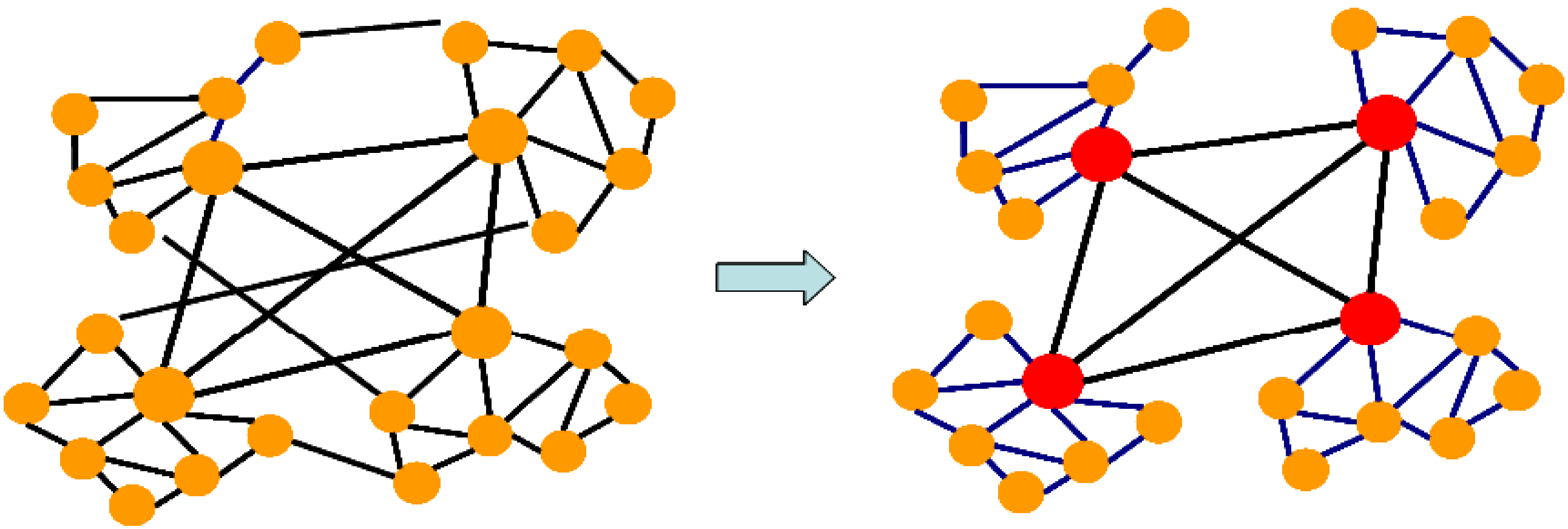
  $N = K \times M, \quad K \leq M \times N$, select 'leaders'

Distributed self - organization

Determine leaders → Find influence → Remove unnecessary links

Semi-decentralized

Decentralized: Dirichlet problem on the graph



Goal: design a scheme that gives each node a vector of compact global information

Two stage semi-decentralized algorithm

- **Stage 1: Determining *K* leaders**
  - Each node determines its social degree via local query
  - Dominant nodes in each neighborhood send their degrees to the central authority
  - Central authority computes their social scores

$$SC(k) = \alpha SD^{(2)}(k) + (1-\alpha) SD^{(3)}(k)$$

    Choice of α determines whether leaders in star-like neighborhoods are preferred

  - The central authority selects the K nodes with highest scores as social leaders and gives them an arbitrary order

Algorithm (cont'd)

- **Stage 2**: **Determining the influence vectors**
  - Based on its order each leader takes its influence vector to be the fixed vector e_i
  - Regular nodes update their influence vector entries:

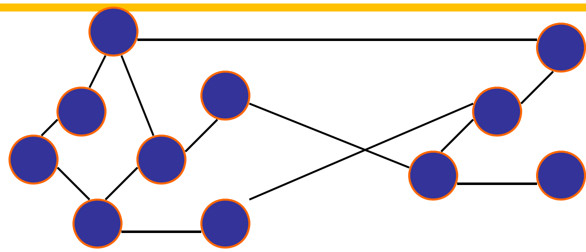
$$x_i^k(t+1) = \frac{1}{n_i + 1}\left[x_i^k(t) + \sum_{j \in N_i(t)} x_j^k(t) \right]$$

- For connected graphs, for t large enough, x_i^k converges to the influence of leader *k* on node *I*

- Upon calculation of influence vectors, each regular node determines its local leader and stops its communication with neighbors who have other leaders

- **Graph decomposes into two level hierarchy with efficient communication pattern**

Reliability and Spanning Trees



$\mathcal{G}(\mathcal{V},\mathcal{E})$

$\mathcal{V} = \{1, 2, ..., n\}$

$\mathcal{E} = \{l_1, l_2, ..., l_e\}$

p : Constant link loss probability

N_i : # of connected components with i edges

$\tau(\mathcal{G})$: Number of spanning trees

$$\mathrm{Rel}(\mathcal{G}, p) = \sum_{i=n-1}^{e} \mathrm{N}_i (1-p)^i p^{e-i}$$

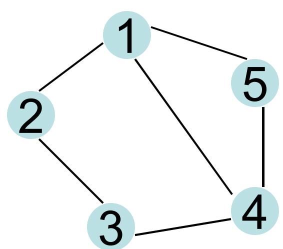
For sufficiently large p:

$$\tau(\mathcal{G})(1-p)^{n-1} p^{e-n+1} \leq \mathrm{Rel}(\mathcal{G}, p) \leq \tau(\mathcal{G})(1-p)^{n-1}$$

- End to end applications
- Spanning tree as a minimally connected graph
- *T(G)* as a measure of robustness to losses
- *References*: Kelmans, Colbourn

Graph Theory for Robust Network Design

- Goal: Given a base topology add *k* edges from a set of *m* candidates such that results in maximum number of spanning trees
- Number of spanning trees $\tau(G) = \frac{1}{n}\prod_{i=2}^{n}\lambda_i(L) = \frac{1}{n}\det(L + \frac{11^T}{n})$
- Incidence vector of an edge shows between which nodes it is



$$f_i : \text{incidence vector}$$
$$f_i = e_\alpha - e_\beta$$
$$f_{(1,5)} = e_1 - e_5 = [1 \quad 0 \quad 0 \quad 0 \quad -1]^T$$

$$f_{15} f_{15}^T = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Graph Laplacian $L = D - A = F_{nm}F_{mn}^T = \sum_{i=1}^{m} f_i f_i^T$

Problem Statement

- Goal: Given a base topology add *k* edges from a set of *m* candidates such that results in maximum number of spanning trees $$\tau(\mathcal{G}) = \frac{1}{n}\prod_{i=2}^{n}\lambda_i(L) = \frac{1}{n}\det(L + \frac{1 1^T}{n})$$
- Dynamic graph process resulting from adding edges

$$\text{Maximize } \tau(G(t+k))$$

Subject to:

$$\begin{cases} G(t+1) = Add(G(t), u(t)), & t = 0,1,\dots,k-1 \\ u(t) = e(t+1), & e(t+1) \in S \subseteq E(\bar{G}(t)) \\ G(t) = \mathcal{G}_0 & \end{cases}$$

Formulation and Relaxation

- Goal: Given a base topology add *k* edges from a set of *m* candidates such that results in maximum number of spanning trees (Approach similar to Ghosh and Boyd 06)

$$\text{Maximize} \quad \tau\left(L_0 + \sum_{i=1}^{m} x_i f_i f_i^T \right) \quad \text{or equivalently}$$

$$\log \det \left(L_0 + \frac{1}{n} J + \sum_{i=1}^{m} x_i f_i f_i^T \right)$$

$$\text{Subject to:}$$

$$1^T x = k$$

$$x \in \{0,1\}^m$$

$$\log \det \left(L_0 + \frac{1}{n} J + \sum_{i=1}^{m} x_i f_i f_i^T \right)$$

is concave in x.

- Relax to $x \geq 0$

$$x_i^* > 0 \Rightarrow \frac{\partial \tau(x^*)}{\partial x_i} \geq \frac{\partial \tau(x^*)}{\partial x_j}, \forall j$$

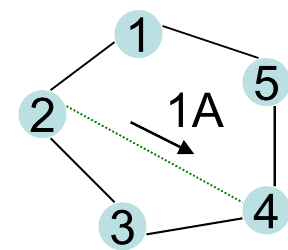
- At maximum $\tau(x)$ has equal derivatives for positive x_i s

Robust Network Design

- Derivative:

$$f_i^T \left(L_0 + \frac{1}{n} J + \sum_{i=1}^{m} x_i f_i f_i^T \right)^{-1} f_i = \lambda, \forall i \in \text{Chosen edge set } (\mathrm{x_i} > 0)$$

$$\mathrm{R_{eff}}(i) = f_i^T \left(L + \frac{1}{n} J \right)^{-1} f_i$$



$$R_{eff}(\alpha, \beta) = V_{\alpha\beta}$$

- If feasible, add edges such that the effective resistance distance of all selected edges become equal and greater than the effective resistance distance between non-selected candidates

Special Cases

- Adding 2 edges (α,β) and (γ,δ)

$$\tau(G(2)) = \Big[\big(1+R_{eff}(\alpha,\beta)\big)\big(1+R_{eff}(\gamma,\delta)\big) - \Big(\underbrace{(z_{\gamma\alpha}-z_{\gamma\beta})}-\underbrace{(z_{\delta\alpha}-z_{\delta\beta})}\Big)^2\Big]\tau(\mathcal{G}_0),$$

$$Z = [z_{ij}] = \left(L + \frac{1}{n}J\right)^{-1}$$

Maximized by adding edge between high resistance distance nodes

Maximized by adding edge to symmetrize the graph

- Adding 3 or more edges similar: more complex terms due to compromising between symmetrizing the graph and joining nodes with the highest resistance distance

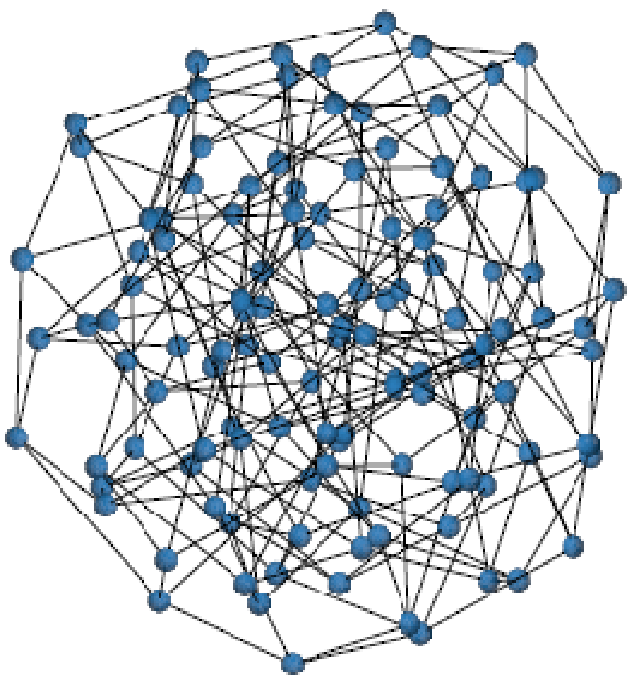
Expander Graphs

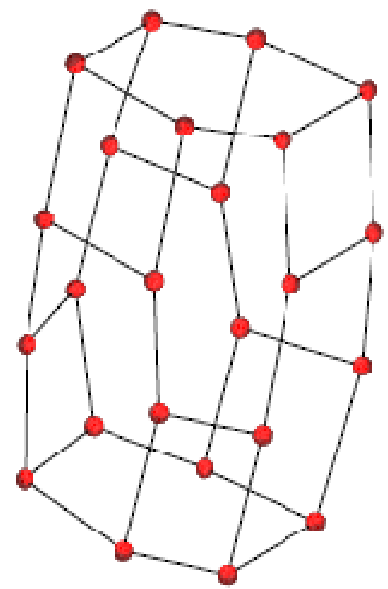
- Fast synchronization of a network of oscillators
- Network where any node is “nearby” any other
- Fast ‘diffusion’ of information in a network
- Fast convergence of consensus
- Decide connectivity with smallest memory
- Random walks converge rapidly
- Easy to construct, even in a distributed way (ZigZag graph product)
- Graph G, ***Cheeger constant h(G)***
  - All partitions of G to S and S^c ,

    $h(G)$=min (#edges connecting S and S^c) /

    (#nodes in smallest of S and S^c)
- (k , N, ε) **expander** : $h(G) > \varepsilon$; **sparse but locally well connected (1-*SLEM*(*G*) increases as $h(G)^2$)**

Expander Graphs – Ramanujan Graphs





Constructing Expander Graphs

- Possible methods:
  - Form a random expander as a $2d$-regular multi-graph in which the set of edges consists of d separate Hamiltonian cycles on APs (Law and Siu 2003)
  - Form a union of two spanning trees chosen independently from the uniform distribution over all spanning trees of a complete graph, implementable by a random walk method (Goyal et al. 2009)

Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

Non-Commutative Probability Models – New Logics

- **Key idea**: interaction between measurements by different agents and between system states and dynamics and measurements (Baras, 1977)
  - Now investigated vigorously in information retrieval systems (van Rijsbergen, 2004)
  - Asynchrony and concurrency
- **Key challenge**: understand the fundamentals of information collection and information flow in multi-agent stochastic control systems
- Witsenhausen's model of information patterns is not correct – even in its most general setting
- Need for new non-commutative probability models – **new logics** --projections in Hilbert space

The Setting and the Problems

- *N* agents, local states, local times
- Measurements and hypotheses supported and interpreted by local states
- Static problems: distributed detection and estimation problems
- Simple dynamics: like in information retrieval systems
- Complex system dynamics: full interactions between measurements and measurements and controls
- Must unify the probabilistic and logical aspects in a consistent way (see recent results of Abbes 2005 for probabilistic models over systems with concurrency constraints)

Models with Incompatibility Build-in

Theorem 5: *If (E, S, P, T) is an event-state-operation structure, then $(E, \leq, ')$ is an ortholattice. Moreover, if $p, q \epsilon E$, then $T_{p \wedge q} = (T_{p'} \tilde{\circ} T_q) \circ T_q$*

Theorem 6: *Assumptions as in Theorem 5 . Then for $p, q \epsilon E$ the following are equivalent:* (a) pCq, (b) $T_p \circ T_q = T_q \circ T_p$. *If pCq then $T_{p \wedge q} = T_p \circ T_q$.*

- **Active interpretation of operations**: can be thought of as a model for the combined operation of taking a measurement and applying a control law by the agent
- **Passive interpretation of operations**: system's interaction to measurements (used by recent results in information retrieval systems)
- We also get an interpretation of the **conjunction of incompatible events or measurements as "data fusion" or "agreement" between agents**

Final Results and Future

Theorem 7: *Databases and conditioning in a multi-agent stochastic control system can be used to construct an event-state-operation structure. This is a faithful representation.*

Theorem 8: *If $(E, \leq)$ is atomic and the atoms $\hat{E}$ are mapped to pure states under T, then (E, S, P, T) can be represented as the lattice of projections on a Hilbert space.*

- The **most useful 'practical' model**

  Finite dimensional Hilbert space, measurements to self-adjoint operators, states to trace one positive operators

- The **biggest payoff**:

  Our theory (extensions of above) allow the formulation of 'design' problems as convex problems over a pair of Banach spaces (one for measurements and one for controls)

- They also results **automatically to introduction of 'supervisors'**

Outline

- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

Biological Networks – Our Research

- Characterization of biological networks
- Discovery of elemental components (e.g. motifs) – Modular decomposition
- Network composition from modules
- Development of a taxonomy of network structure *vs* behavior *vs* biology
- Network dynamics and their interpretation
- Network inference and tomography
- Applications to disease pathology (e.g. cancer)
- Development of analytic/computational tools

Cancer and Systems Biology -- networks are key

- Emergent properties at the system level – not just at components
- Multiple-interacting networks – gene networks, protein networks, metabolic networks
- Network dynamics more important
- Complex diseases cause changes in network dynamics
- **Key Question**: detect these changes from repeated partial measurements (data) at different scales

Network Inference from Data

- Inference of network models, from data and prior structure knowledge
- Various types of dynamic networks: Networks of ODEs, Bayesian Networks, Boolean Networks, Hybrid Networks
- Stochastic Graph Processes, their representation and parameter/structure estimation, Time Varying MRF
- **Usage**: Network level analysis to improve cancer prognosis (e.g. metastasis in breast cancer via sub-networks of the protein-protein interaction network)
- Use alterations of the molecular network in malignant cells to identify oncogenes (e.g. for B-cell lymphomas)
  - Common alteration: upregulation of growth factor receptors (e.g. EGFR hyperactivated in several cancers

Use of Network Models and Associated Analysis in Cancer

- Develop multi-network models and use them to understand and detect crucial characteristics of cancerous cells:
- Independence from external growth signaling
- Insensitivity to antigrowth signaling and evasion of apoptosis
- Limitless replicative potential
- Sustained angiogenesis and metastasis
- **Example 1**: Use a human protein-protein interaction network model -- compute certain sub-networks – prove that they re good indicators of metastasis (used maximum mutual information and conditional likelihood classification) in breast cancer [Chuang et al, 2007]
- **Example 2**: Develop a network centric approach to distinguish normal from malignant cells and identify targets and effectors of specific biochemical perturbations , potentially useful for the identification of drug targets [Mani et al 2008]

Example: Dynamic Modularity in Protein Interaction Networks for Breast Cancer Prognosis

- [Taylor et al , 2009]
- Dynamic structure of human interactome can be used to predict patient outcome
- Identify inter-modular hub proteins co-expressed with their interacting partners and intra-modular hub proteins co-expressed with their interacting patterns
- Observed substantial differences in biochemical structure of two hub types
- Signaling domains found in inter-modular hub proteins – associated with oncogenesis
- Analysis with breast cancer patients – altered interactome useful as indicator of breast cancer prognosis

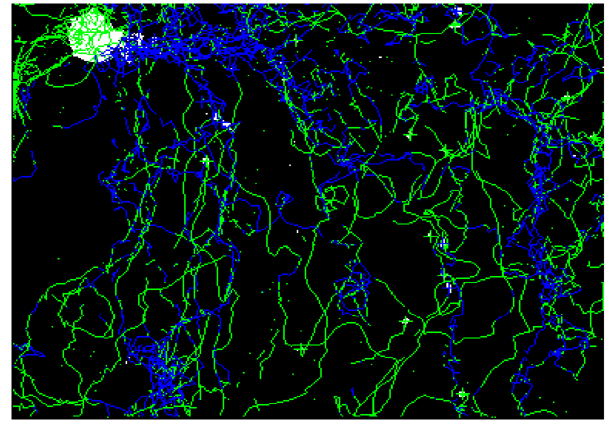
Outline

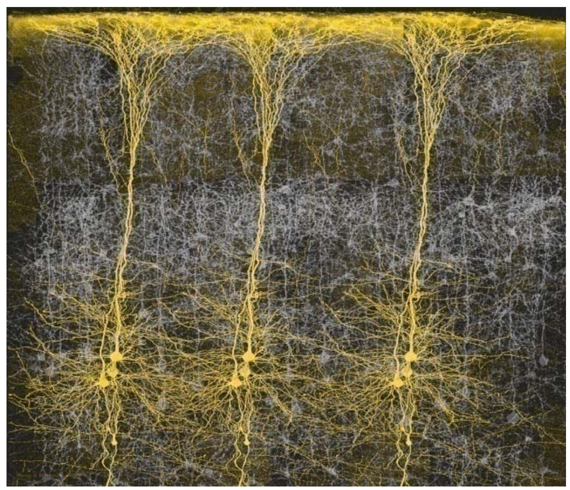
- Multiple interacting dynamic hypergraphs – four challenges
- Networks and Collaboration -- Constrained Coalitional Games
- Trust and Networks
- Component-based network synthesis
- Topology and performance
- New probability models (non Kolmogorov)
- Biological networks and cancer dynamics
- Conclusions and Future Directions

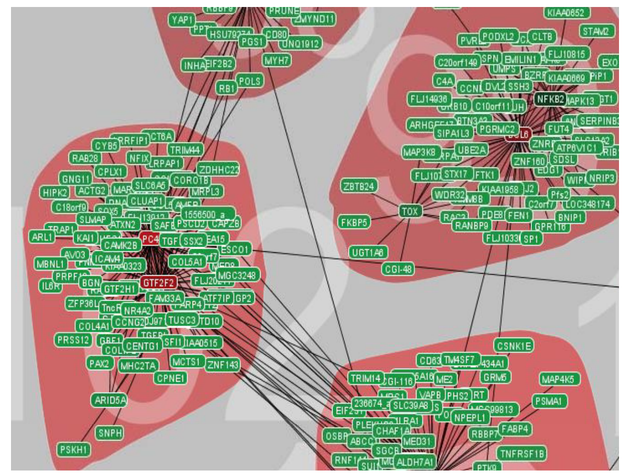
Conclusions

- Complex networks – multiple dynamic hypergraphs
- Fundamental tradeoff between the benefit from collaboration and the required cost for collaboration
- Coalitional games and network formation
- Trust as a catalyst for collaborations
- Component based network synthesis
- Effects of topology on distributed algorithm performance
- Performance vs. efficiency – small world graphs – expander graphs
- New probabilistic models (non-Kolmogorov)
- Biological networks and cancer dynamics

How Biology Does IT?







Lessons Learned -- Future Directions

- **Constrained coalitional games** – unifying concept
- Generalized networks, **flows - potentials**, duality and network optimization (monotropic optimization)
- **Time varying graphs** – mixing – statistical physics
- **Understand autonomy** – better to have self-organized topology capable of supporting (scalable, fast) a rich set of distributed algorithms (small world graphs, expander graphs) than optimized topology
- Given a set of distributed computations **is there a small set of simple rules** that when given to the nodes they can self-generate such topologies?

Thank you!

baras@isr.umd.edu

301-405-6606

http://www.isr.umd.edu/~baras

Questions?