Learning for Heterogeneous Networks: Hierarchical Representation

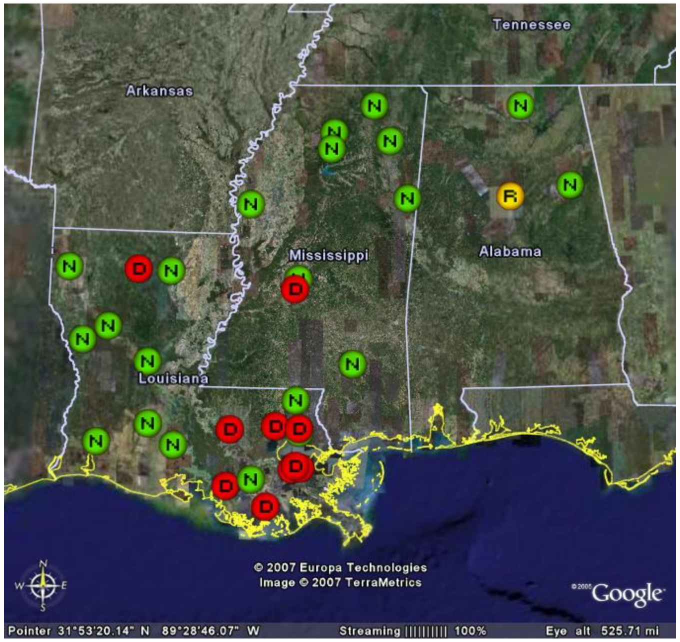
Chuanyi Ji
ECE, Georgia Tech

Acknowledgement: Based on joint work with S. Jeon, G. Liu, S. Erjongmanee

Network and Natural Disaster

Natural disaster: Hurricane Katrina 2005, Ike 2008, Gustav 2008…

Inferring network service disruption: 1000+ subnets/variables, 40+ ISPs, 33% disruption: > 1 hour

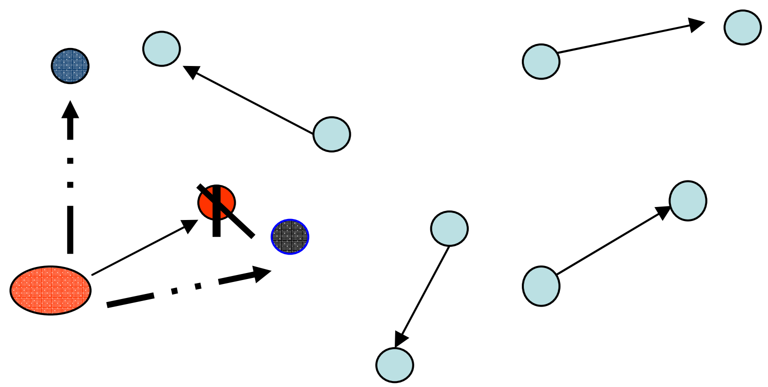


S. Erjongmanee and C. Ji, "Inferring Network Service Disruption Upon Hurricane Katrina: Large-Scale Sensory Measurements and Human Inputs," Sensor-KDD 08

Network and Natural Disaster

- Goal: understand causes/vulnerability
  - Physical damages, power shortages, evacuation, …
- Hard to infer dependence:
  - Physical failures ⟷ Measurements?
  - Disparate data?

Wireless Ad-hoc Network: Self-Configuration



Upon failure: Nodes move, re-establish transmission

Configuration: physical topology, transmission (channel reuse), number of nodes, density, traffic, channel…

What local decisions result in an optimal configuration?

Question

- How does network performance (resilience and optimal configuration) depend on many variables, external and internal?

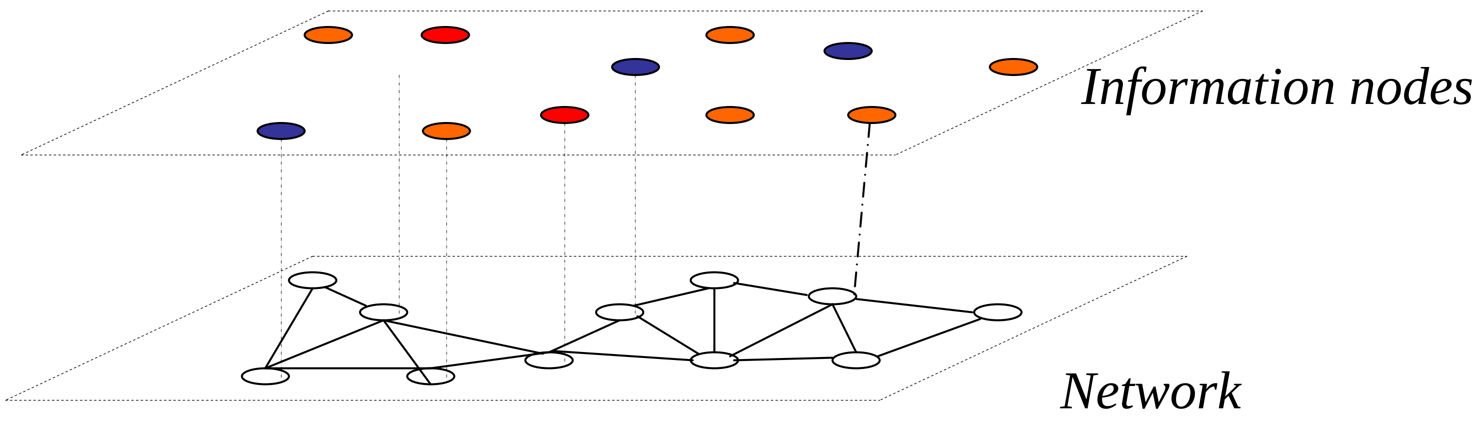
Outline

- Representation of Dependency
  - Mapping from networks
  - Performance and complexity
    - Is distributed operation optimal?
  - Open Issues

Variables

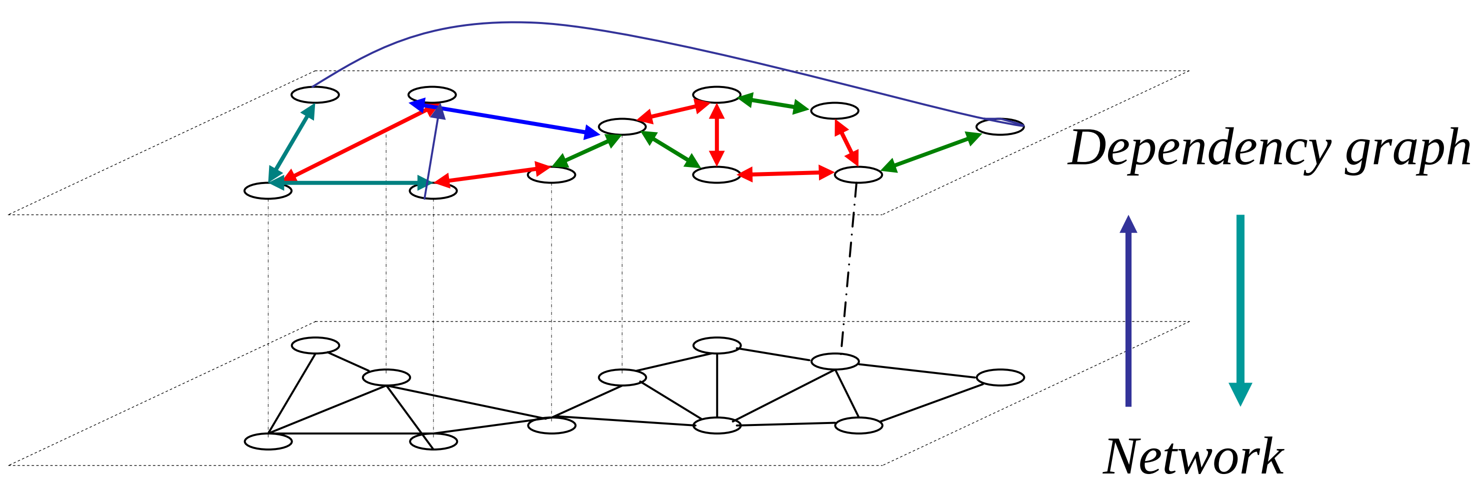
- Information nodes: $Z = \{Z_i\}_{i=1}^{n}$

  Z_i : nodal states, aggregated flows, positions, events, decisions



Problem Statement

- Dependency graph: G(Z, E), E: Links (dependency)
- Learning representation:
  - Find G(Z,E) given a network architecture



Prior Work

- Probabilistic Graphical Models:
  - Markov Random Field (Geman84): Undirected, small-world for images
  - Probabilistic graph (Jordan92): Directed, undirected.
  - Factor graph (Kschischang01): General framework and message passing

Many applications in image processing, biology, …

Prior Network Applications

- Model-based:
  - Simple graphs for analyzing network performance [Mitra93][Zachary96][Niu06]
  - Homogeneous variables

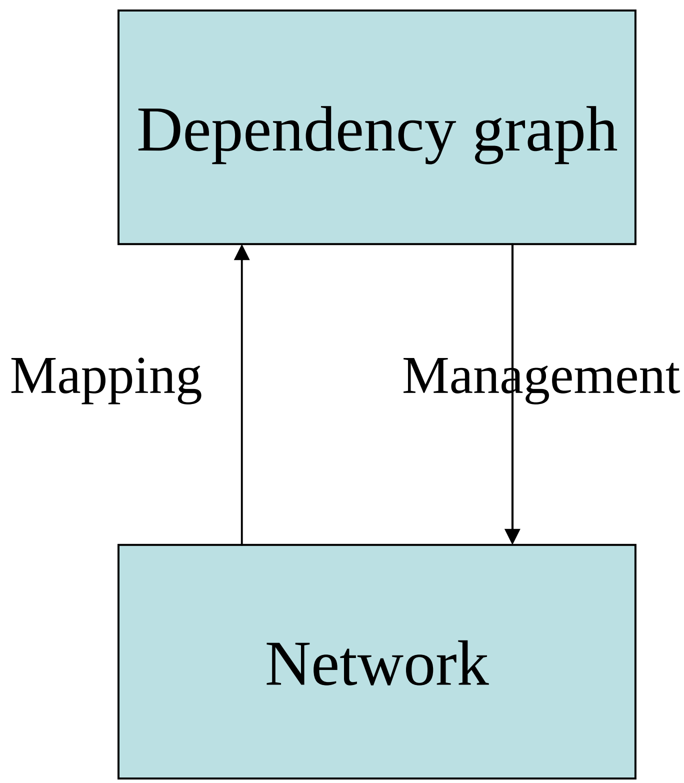
– Data driven:
  - Learning dependence from network data [Katebi07][Greenberg07]…

– Applications still developing

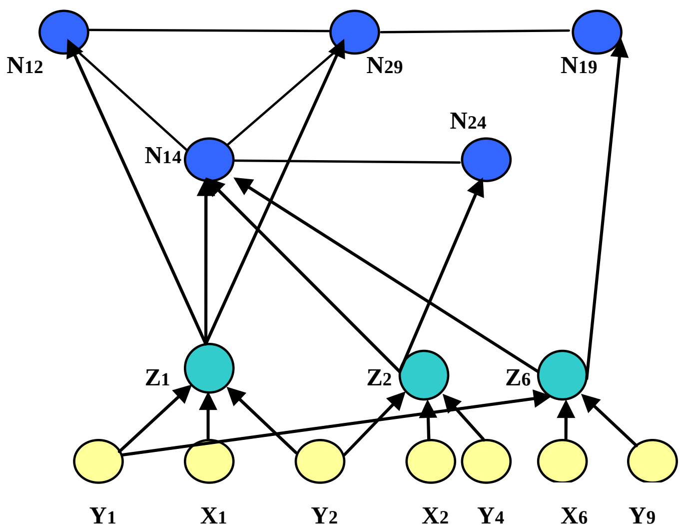
Our Approach

- Initial idea: Apply graphical models directly to network data

- Bottom-up approach: Network determines dependency graph



Example 1: IP Network

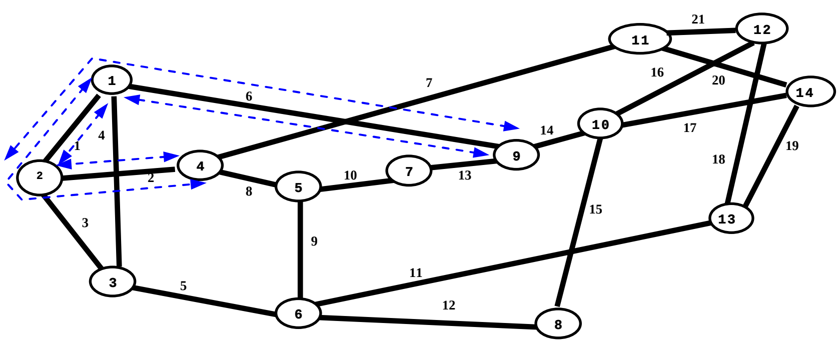


Dependency graph:

Flow: N_{ij}

Link status: Z_i

Failure events: X_i, Y_j



Network:

Given topology, link capacity, random flows

G. Liu and C. Ji, "Network Resilience: Multi-Layer Graphical Model and Analysis of Scalability," to appear IEEE/ACM Trans. Networking

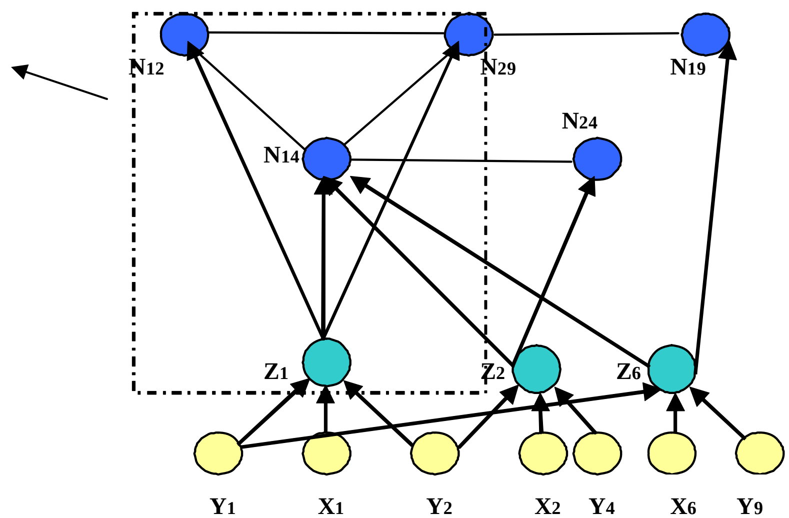
Mathematical Representation

$$P(Z) = \prod_{i=1}^{n} \Phi_i(Z_i, Z_j, \forall j \in C_i)$$ (Geman, Jordon, Kschichang)

$\Phi_i()$: Clique function of Z_i , C_i : Neighborhood of node *i*

Conditional independence: $P(Z_i \mid Z \setminus Z_i) = P(Z_i \mid Z_j, \forall j \in C_i)$

Clique for N_{12}



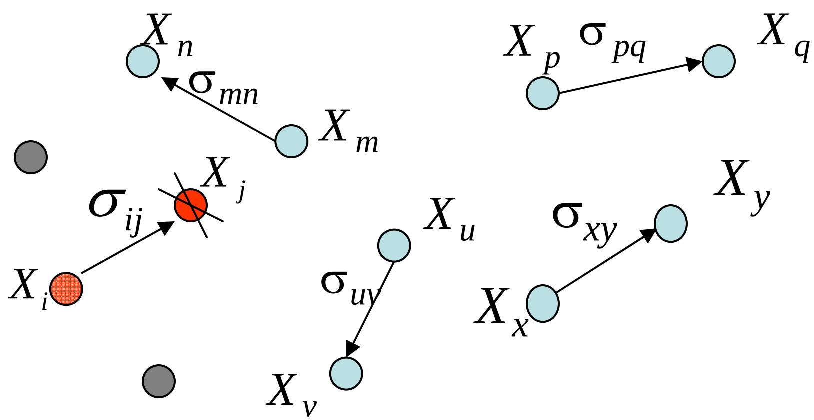
Analysis of Network Resilience

Resilience: % retained traffic due to physical failures P(Z)

Table 1: Network Resilience

Resilience	$\rho \to 0$ (low load)	$\rho \to 1$ (high load)
Ring network	$\rho(3+\alpha)/(m-1)$	$O(1/m)$
Star network	$\rho(3+\alpha)/m$	$O(\alpha)$
Mesh-torus	$2\rho(3+\alpha)/(m-1)$	$\begin{cases} O(1/(1-\alpha)m), & if\ \alpha \neq 1, \\ O(1/\sqrt{m}), & otherwise. \end{cases}$

Example 2: Wireless Ad-hoc Network



Random position: X_i

Transmission: σ_{ij}

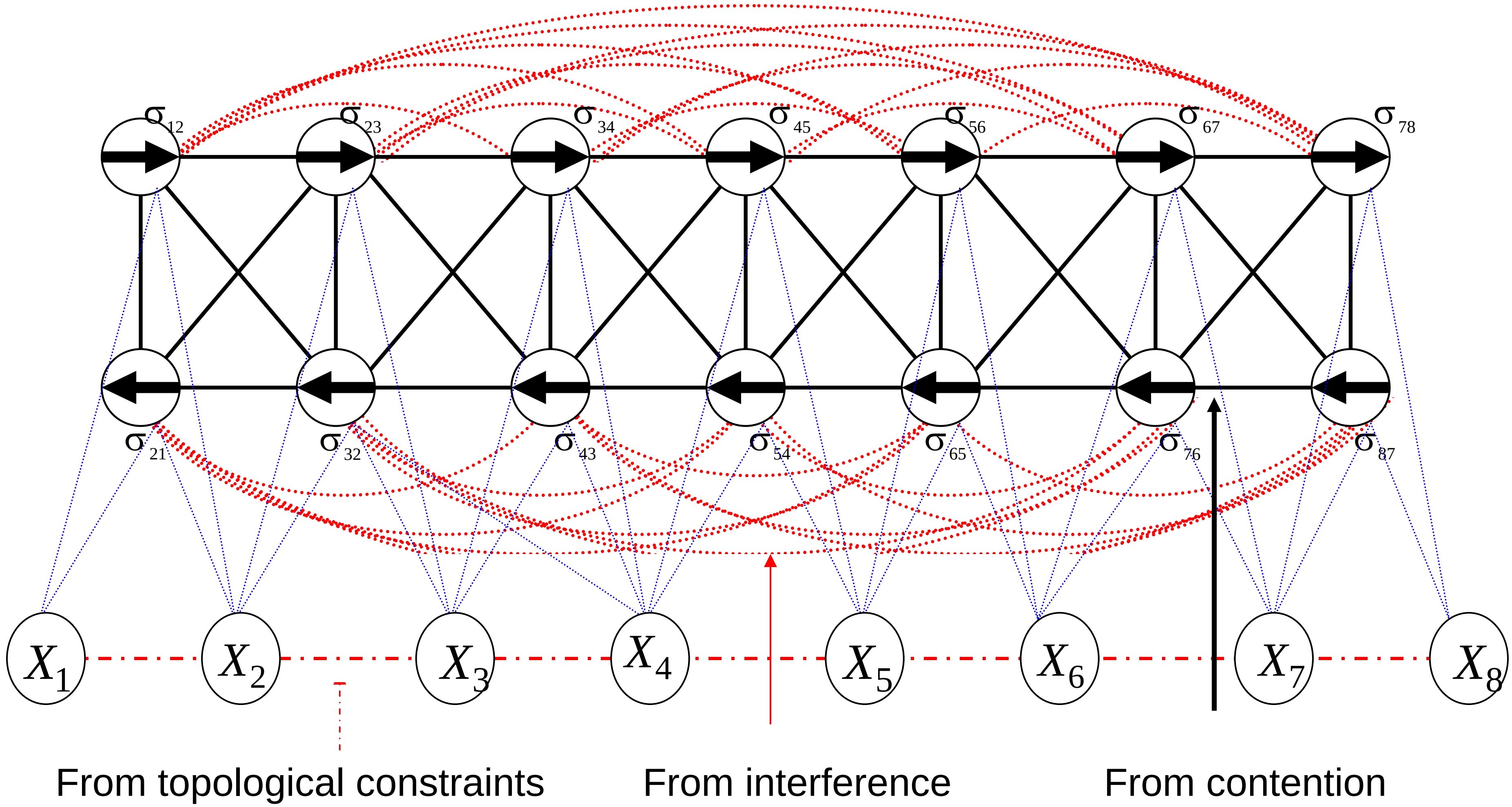
Node decision: $Z_i = (X_i, \sigma_{ij})$

Assumptions [Jeon05,08]:

- One carrier frequency, omni-directional antenna, path-loss $\alpha \geq 2$
- Node *i* can communicate with node *j* in a transmission range
- Management requirements: SINR, 1-connectivity, cost
- Only consider physical and link layers.

S. Jeon and C. Ji, "Near-Optimal Distributed Configuration Management of Ad Hoc Wireless Networks Using Probabilistic Graphical Models,'' Lecture Note Computer Science, 2005, axiv 2008

Dependency Graph: Random Bond

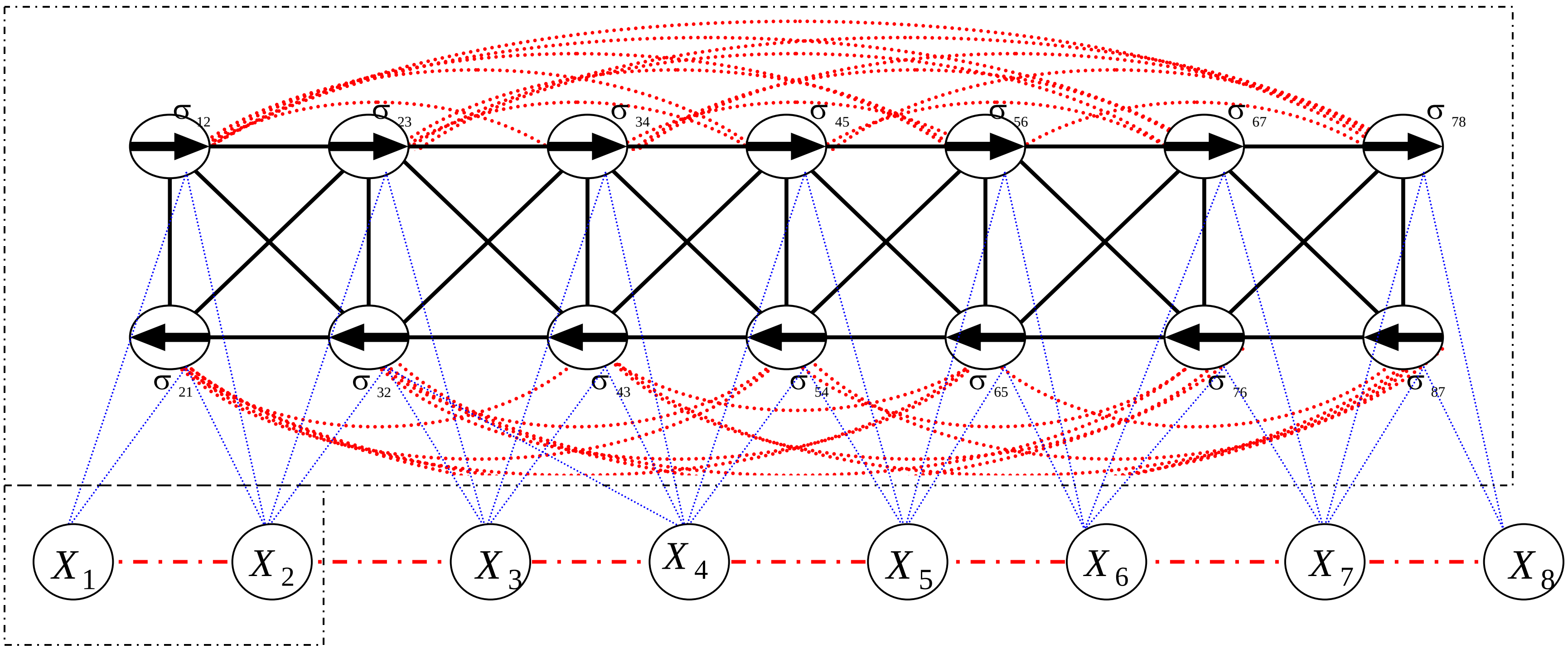


Mathematical Representation

$$P(Z) = \prod_{i=1}^{n} \Phi_i (Z_i, Z_j, \forall j \in C_i)$$

P(Z)=B exp[-f(Z)], Gibbs distribution

Clique for σ_{12}



Commonality

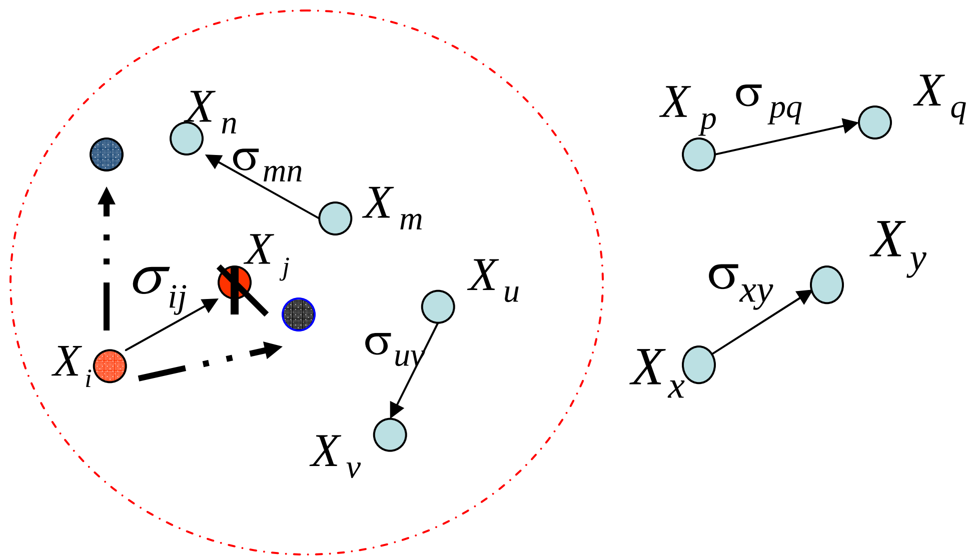
- Hierarchical Graph Structure
  - Heterogeneous variables at different spatial scales
  - One layer: homogeneous variables
  - Dependence within and across layers
- Mathematical representation
  - P(), decomposable to clique functions

Difference

- Variables:
  - Different sources of randomness
- Locality:
  - Clique-size, pattern

- Representation shows:
  - The dependencies are not exactly local
    - Long flows in wireline networks
    - “Long-range” Interference in wireless networks
  - When is distributed self-configuration near-optimal?

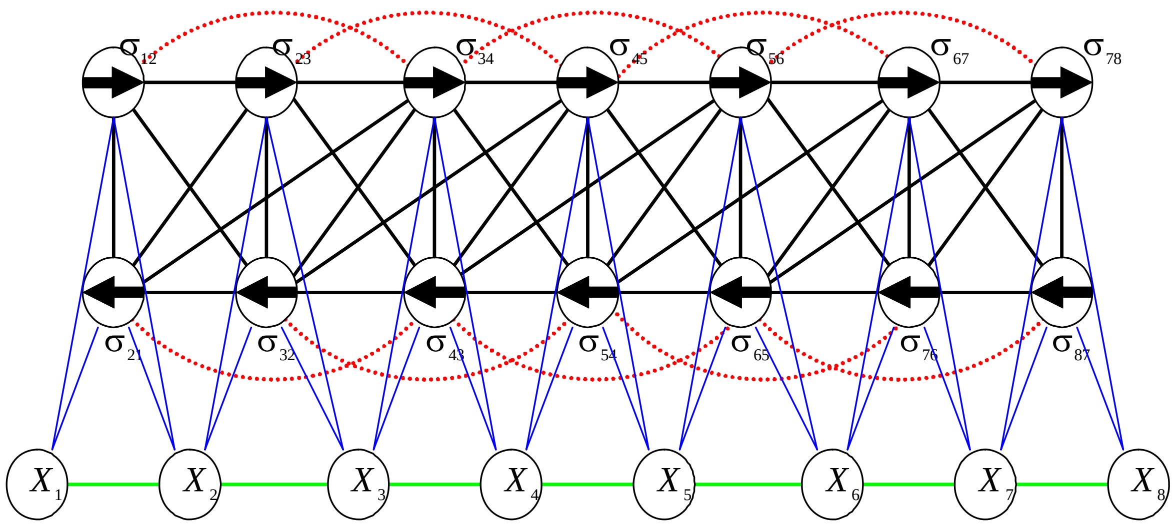
Example 2: Wireless Adhoc Network



Nodes make local decisions on position and transmission using neighbor information:

Determine X_i, σ_{ik} using $P(X_i(t), \sigma_{ik}(t) \mid \sigma_{mn}, X_m, m \in C_i)$

Approximated Dependency Graph



Neglecting long-range interference

Approximation

P() is approximated by two-layer Markov Random Field

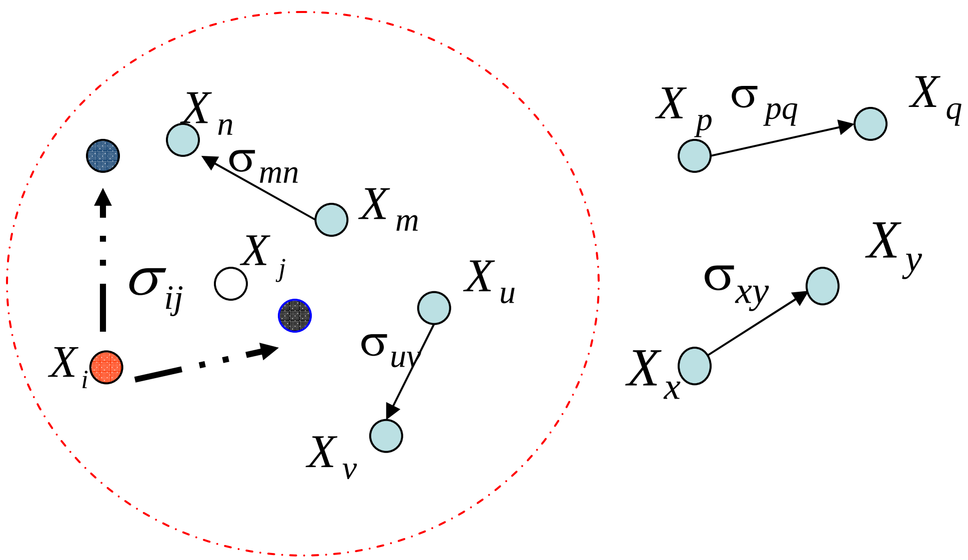
$$P^l(\sigma, X) = \prod_i P_i(\sigma_{mn}, X_m, m \in C_i)$$

$P_i(\sigma_{mn}, X_m, m \in C_i)$: Local likelihood

Performance

- Ground truth: $P(Z^*),\ Z^* = (X^*, \sigma^*) : \max P(Z)$
- Approximation: $P^l(\tilde{Z}),\ \tilde{Z} : \max P^l(\tilde{Z})$
- Performance: $$E\left[\frac{\log P(Z^*) - \log P\left(\tilde{Z}\right)}{\log P(Z^*)}\right]$$
- Complexity: Neighborhood size

When is Local Decision Near-Optimal?



Intuition: Near-optimal if

Sum of neglected dependence < decay of dependence

neighborhood size, density

Channel

Sufficient Condition for Near-Optimality

A uniform network, *n* nodes
Channel: power decay α
Desired performance: $E() \leq \varepsilon$
Complexity (neighborhood size): n_c

$$n_c \geq \begin{cases} \Omega(n^{\frac{4-\alpha}{4+\alpha}}), & 2 \leq \alpha < 4, \\ \Omega(1), & \alpha > 4. \end{cases}$$

Open

- General theory?
  - Hierarchical representation:
    - From Hour glass?
    - What granularity should information be?
    - Relation to optimization (Geman92, Chiang et.al 07, Xu07)?
  - Locality:
    - Flows?
    - Multiple ISP networks?

Open

- Use of representation
  - Help learning large data sets?
- Temporal dependence?

Acknowledgement

- Co-authors: S. Erjongmanee, S. Jeon, G. Liu
- Discussions with A. Walid, W. Willinger, M. Chiang
- IPAM organizers