Network Science and Its Application in Wireless Networking

Dapeng Oliver Wu
Electrical & Computer Engineering
University of Florida

Outline

- Definition of network science
- Major problems in network science
- Application domains
  - Power grid
  - Cognitive/logical/semantic networks
  - Economic networks, e.g., stock markets
  - Biological networks, e.g., brain networks
  - Social networks
  - Information networks
  - Hybrid: information network & social network
- Major approaches in network science
- Joint congestion control and scheduling in wireless networks

What is Network Science? (1)

- Network science is a new and emerging scientific discipline concerned with the interconnections among elements within a system/network.
  - physical or engineered networks (e.g., power grid, and transportation network)
  - information networks
  - biological networks (e.g., brain network, gene regulatory network, neural network)
  - Cognitive/logical/semantic networks (e.g., concept/topic networks)
  - economic networks (e.g., stock market)
  - social networks

What is Network Science? (2)

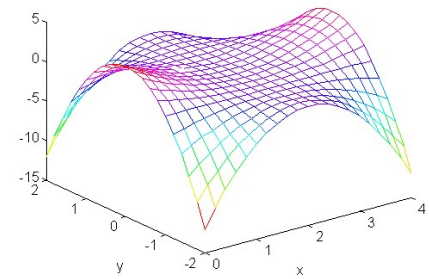
- This field of science seeks to discover common principles, algorithms and tools that govern
  - network structures/topologies
  - network functionalities
  - network behaviors.

Three Major Problems in Network Science

- Network sampling problem:
  - network Nyquist sampling theorem?
- Graph learning problem:
  - identify the network structure/topology.
- Identification of network dynamics/behaviors:
  - identify the ODE/PDE/SDE model of each node in the network
  - estimate the parameters of the model

Network Sampling Problem

Given a 2D random field $f(x, y)$, where (x, y) can be continuous or discrete, find the optimal locations of N sensors with measurements f_i $(i = 1, \cdots, N)$ such that MSE $\iint (f(x, y) - \hat{f}(x, y))^2 dxdy$ is minimized, where $\hat{f}(x, y)$ is reconstructed from f_i $(i = 1, \cdots, N)$.



- Does Network Nyquist sampling rate exist?
- Solving this problem has broad implications, e.g.,
  - How to deploy sensors so that the probability of detecting vehicles/tanks, gunshots, fire in the area is maximized
  - How to deploy traffic monitors so that the probability of detecting the locations of anomalies in communication networks is maximized
  - How to deploy sensors so that the probability of detecting terrorist activities is maximized

Graph Learning Problem

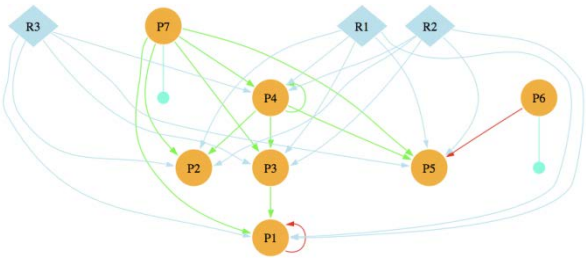
For a network of K nodes, assume the state of each node can be measured.

Given N samples of the state of each node $x_{i,j}$, ($i = 1,\ldots,K;\ j = 1,\ldots,N$)

where i denotes node index, and j denotes sample index,

identify the network topology, i.e., whether there is a link between

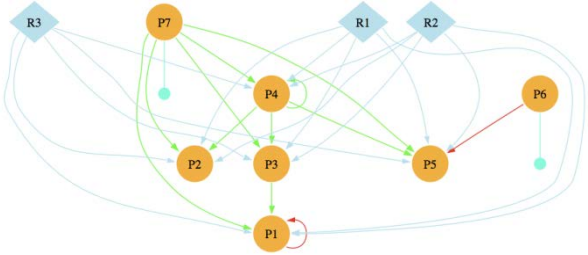
Node i and Node k, where $i \neq k$, and $i,k = 1,\ldots,K$.



- Implications:
  - Identify the topology of
    - gene regulatory networks, given micro-array data
    - social networks, given telephone records of a community, or messages in Internet chat-rooms.
    - semantic networks of words/phrases, given text documents
    - brain networks/connectome, given fMRI or Brainbow images

Identification of Network Dynamics/Behaviors

- Assume the state of each node in a network is time-varying
  - Model each node by a dynamic system
  - Characterize each node by ODE/PDE/SDE
  - Problem: given samples of the state of each node, identify/estimate the parameters of ODE/PDE/SDE
- Implications:
  - A gene regulatory network may change over time; need to model it as a dynamic network
  - Same for a social network



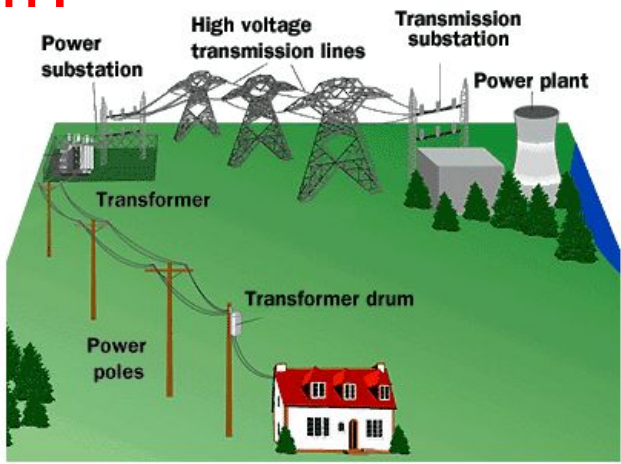
Outline

- Definition of network science
- Major problems in network science
- Application domains
  - Power grid
  - Cognitive/logical/semantic networks
  - Economic networks, e.g., stock markets
  - Biological networks, e.g., brain networks
  - Social networks
  - Information networks
  - Hybrid: commodity/energy/money/human flows
- Major approaches in network science
- Joint congestion control and scheduling in wireless networks

Power Grid

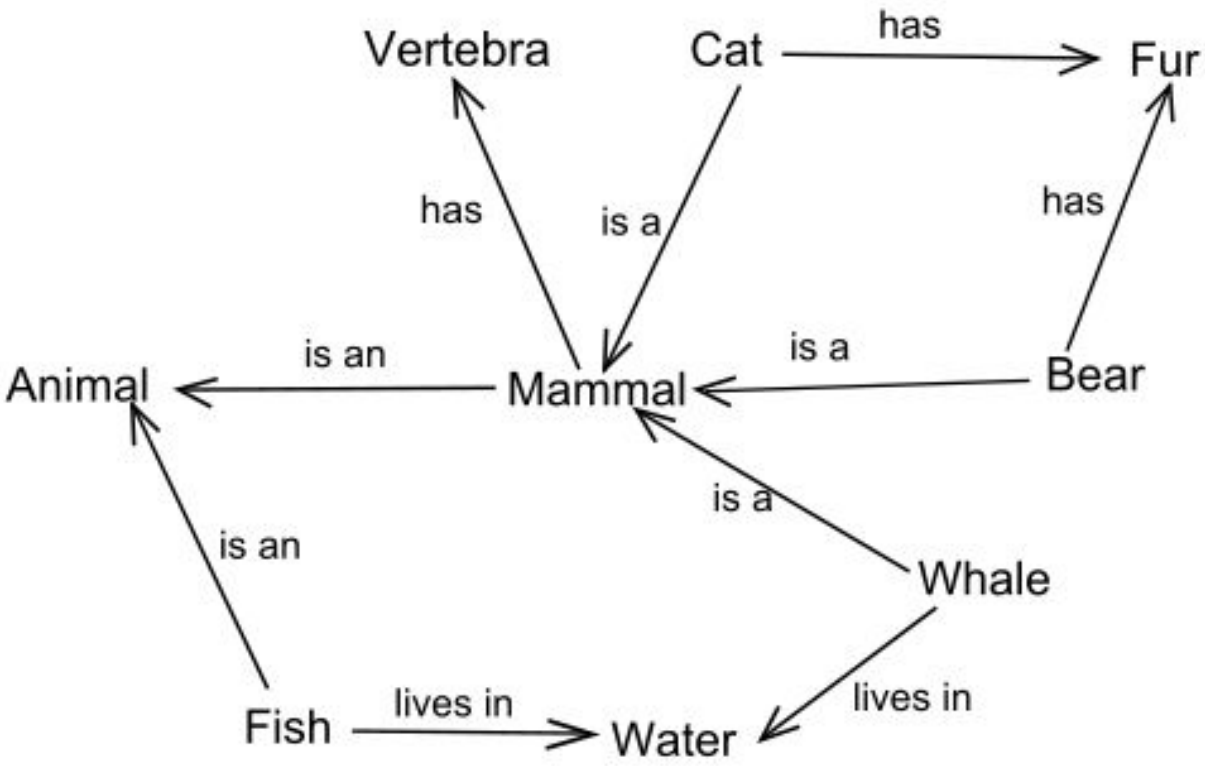
- How does network topology affect the robustness? Vulnerability analysis
- How to model the dynamics of large-scale power grid?
- Is it possible to predict large-scale blackouts?
- How to prevent large-scale blackouts?

⇨ very challenging problem!



Semantic Networks

- How to extract a semantic network, given samples such as terms and text documents



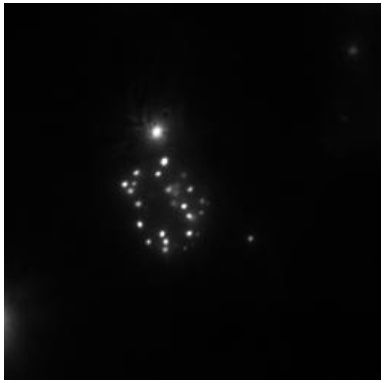
Economic Networks

- How to detect bubbles in stock market, given stock prices and news

  ⇨ you can make a fortune out of it, if you can predict a bubble accurately
- How to detect economic bubbles, given economic data

Biological Networks

- Problem: identify those nuclear bodies whose behaviors are correlated, given fluorescence microscopy images
- Implications
  - Physical bio-marker for early cancer detection
  - Help understand the mechanisms of cell movement (motility)

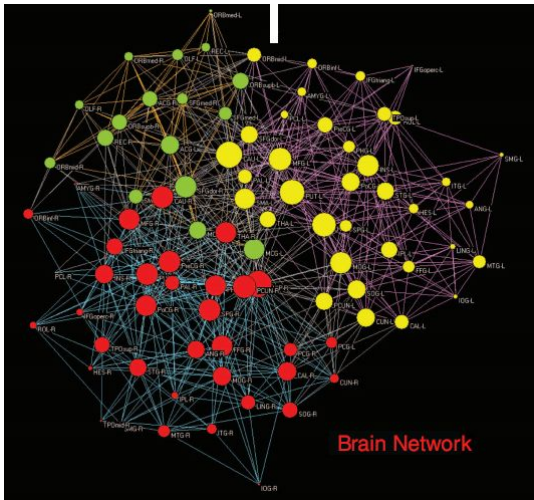


Cancer cell

Normal cell

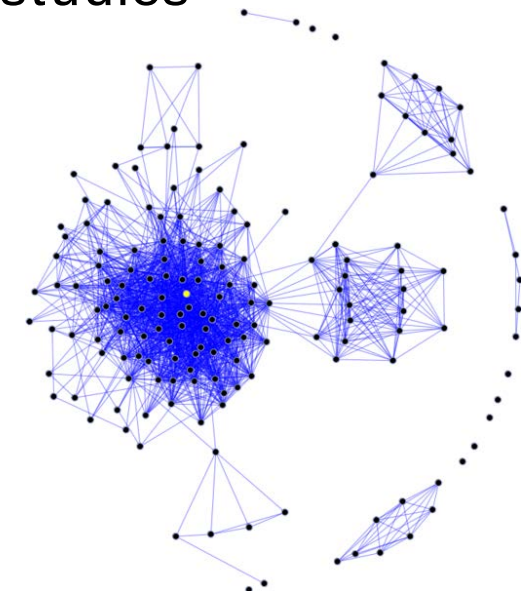
Brain Networks

- How to identify the structure of a brain network, i.e., brain connectome
- Parcellation/clustering, registration, change detection
- Implications
  - A connectome is fundamentally important for understanding brain growth, aging, and abnormity
  - Abnormal brain connections may indicate schizophrenia, multiple sclerosis, autism, or Alzheimer's disease



Social Networks

- How patterns of human contact affect certain phenomena, e.g., the spread of diseases, cultural changes, happiness?
- Implications
  - anthropology, epidemiology, biology, communication studies, economics, geography, information science, organizational studies, social psychology, sociology, sociolinguistics, political science, cultural studies



Information Networks

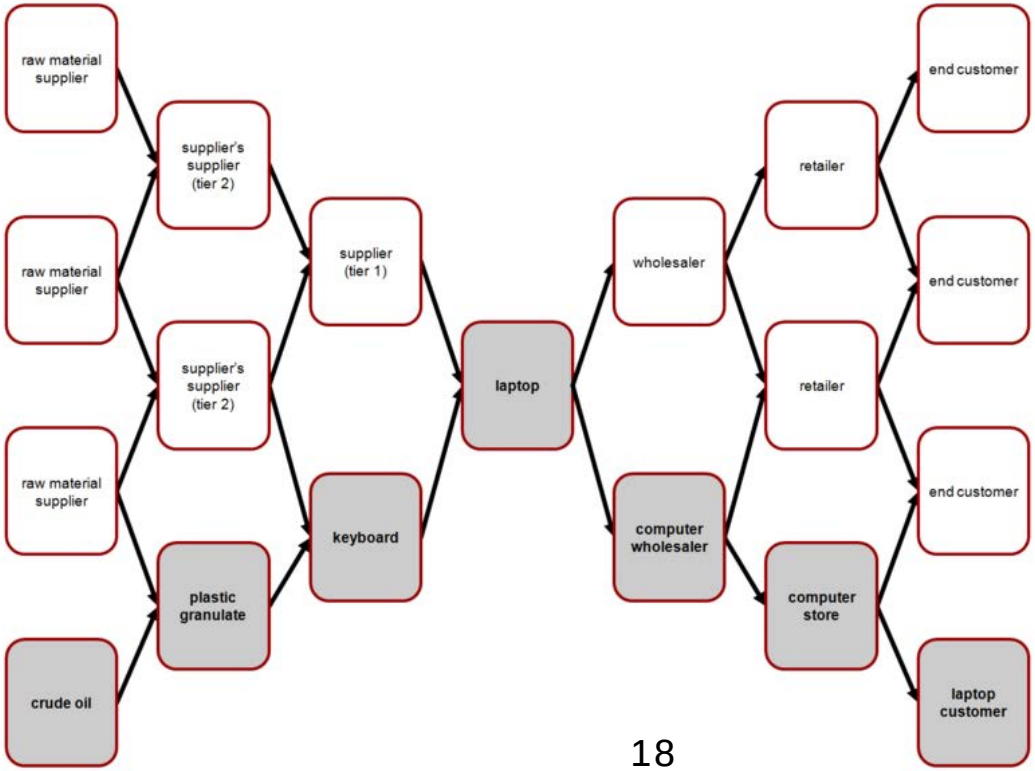
- What is the interplay between communication networks and the social networks involved
  - Online communities, e.g., Twitter, Facebook, LinkedIn
  - Wired/wireless networks, iPhone, laptop, desktop
  - Location of users: office, café, subway, bus
- Implications
  - The insights gained may help design future communication networks with desired robustness, efficiency, and security

New Theory for Traffic Engineering

- Erlang-B and Erlang-C formulae are no longer accurate for scale-free telecom/computer networks, due to:
  - the call frequency of the phone numbers called by a caller follows power law
  - the holding time follows a heavy-tail distribution instead of an exponential distribution, which was assumed in Erlang-B and Erlang-C formulas.
  - the topology of the telecom/computer network is scale free
- A new theory for traffic engineering is needed.

Hybrid Networks

- E.g.: Networks in economy, supply chain management, ecology
- commodity/energy/money/human flows



Major Approaches in Network Science

- Boltzmann principle
- Gibbs principle (decomposition)
- Message-passing method
- Approximation algorithms for NP hard problems
- Compressive sensing approach for NP hard problems

Boltzmann Principle

- Study asymptotics of a large scale problem:
  - It may be hard to obtain a solution for a small-scale problem. But a large scale problem may admit a simple solution.
  - Convert a combinatorial problem into a much simpler problem
  - E.g., Asymptotic Equipartition Property (AEP), typical set, channel capacity, large deviation theory

Gibbs Principle

- Gibbs principle (decomposition):
  - Decompose a high-dimensional problem into multiple low-dimensional sub-problems
  - E.g.,
    - Gibbs sampler
    - Expectation-Maximization (EM) algorithm
    - Blahut algorithm
    - Dynamic programming, Viterbi algorithm

Message-passing Method

- Design a (near) minimum-complexity algorithm that leverages the network structure to recursively solve a large-scale problem.
- E.g., belief propagation algorithm for inference on graphical models

Approximation Algorithms for NP Hard Problems

- Exact algorithm:
  - computationally infeasible in practice
- Approximation algorithm:
  - achieves theoretically proven good performance, e.g., close to the optimal
  - Polynomial time but may have high complexity
- Heuristic algorithm:
  - lower complexity than approximation algorithm but no performance guarantee

Compressive Sensing Approach for NP Hard Problems (1)

- Compressive sensing (CS) is a technique for acquiring and reconstructing a signal utilizing the prior knowledge that it is sparse or compressible.
- A new paradigm to avoid NP hardness
  - Convert a mixed-integer program to a nonlinear program
  - Solve the nonlinear program
  - Decoding: map the real-valued solution to integer valued feasible solution (a valid codeword)

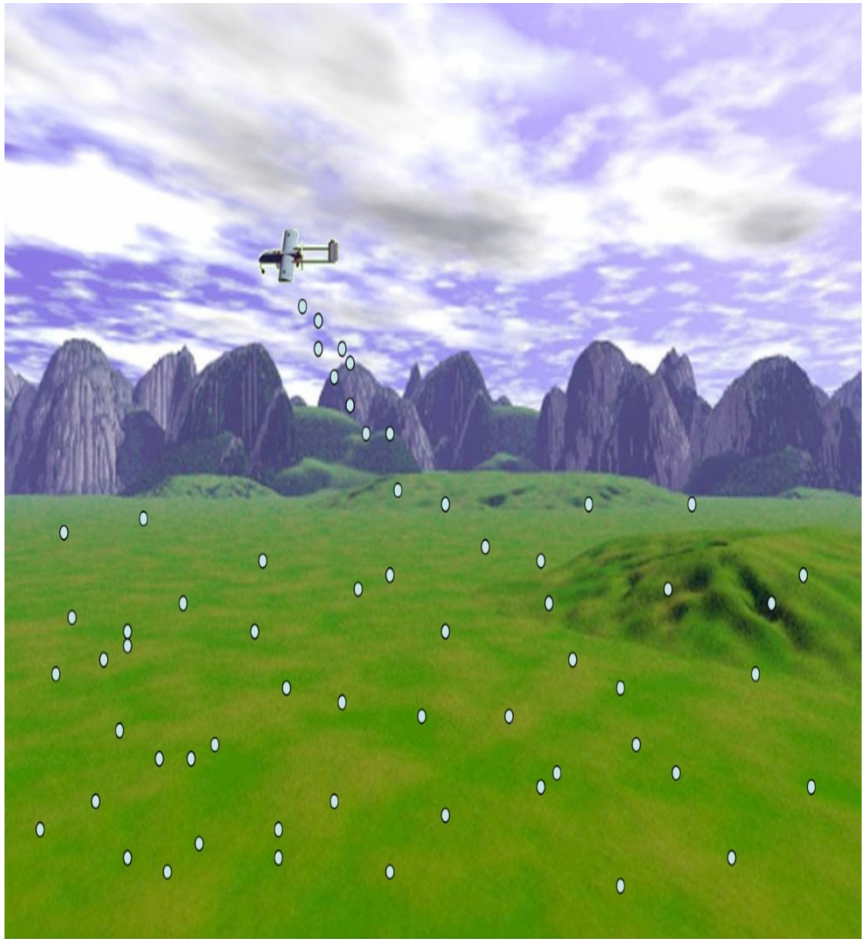
Compressive Sensing Approach for NP Hard Problems (2)

- CS = Model selection = penalized least squares
- We can apply CS to
  - the graph learning problem
  - parameter estimation problem of a large-scale network

Joint Congestion Control and Scheduling in Ad Hoc Networks

Using Decomposition Principle

Optimization in Ad-hoc network



Challenges in Ad-hoc:

Resource Allocation :

- Congestion Control
- Scheduling
- Efficiency/Fairness

QoS awareness

- Elastic (email)
- Inelastic (voice)

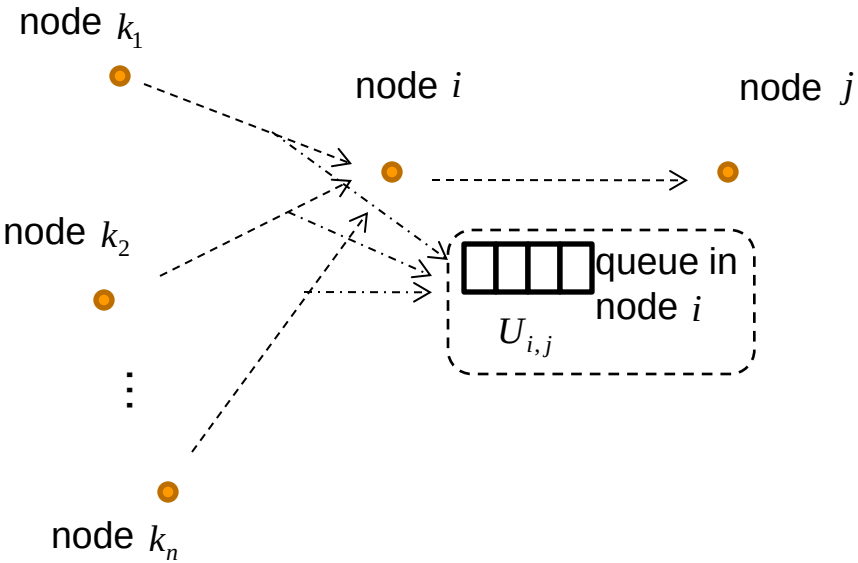
Decentralized algorithm needed:

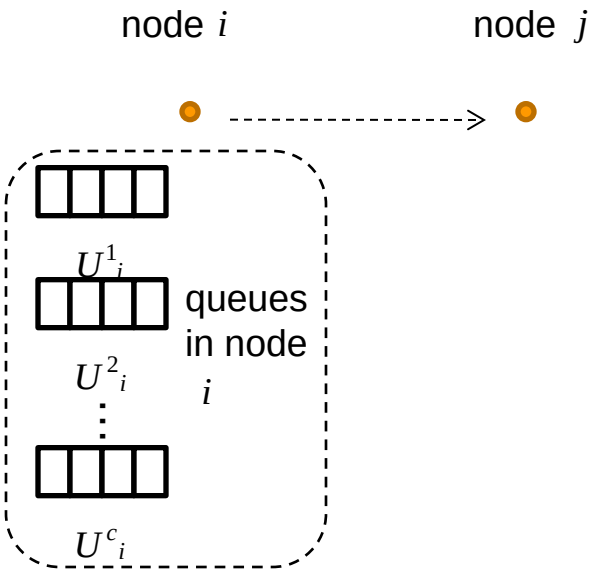
- Less overhead
- Fast convergence

Problems and Solutions

- Existing works:
  - Network Utility Maximization (NUM) problem[Kelly98]
    - Maximum Weight Scheduling [Tassiulas92]
    - Dynamic Routing and Power Control (DRPC) [Neely, 03]
  - Capacity region for per-destination queuing system [Neely03]
  - Decomposition between congestion control and scheduling [Lin04] [Chiang07]
- Our Solutions:
    - per-next-hop queuing system
      - Reducing queuing overhead
      - Extending the definition of conventional capacity region
    - Sliding Mode (SM) based controller in wireless network
      - Providing distributed iterative algorithm to optimization problem with a large number of constraints [Utkin92] [Lagoa04]

Per-next-hop queues





Per-destination queue (PDQ)

- Defined on nodes
- The number of queues in each node equals to the number of destinations

Per-next-hop queue (PNHQ)

- Defined on links
- The number of queues in each node equals to the number of neighbors

Problem Formulation

- The NUM problem in per-next-hop queuing system

$$\max_{\{\mathbf{x},P\}} \sum_{s\in\Theta^f} U_s(x^s) \tag{5a}$$

$$s.t.: \quad \mathbf{x} \in Co(\mathcal{R}) \tag{5b}$$

$$\sum_{s:b_s=i} x^s_{i,j} - R_{out(i,j)} + R_{in(i,j)} \leq 0, \forall\ i \in \Theta^n, j \in\ S_i \tag{5c}$$

$$P \in \mathcal{P} \tag{5d}$$

$$\mathbf{h} \leq 0 \tag{5e}$$

- The dual decomposition w.r.t. the flow conservation constraint

$$\begin{aligned} D(\Lambda) &= \max_{\{\mathbf{x},P\}} \sum_{s\in\Theta^f} U_s(x^s) - \sum_{i\in\Theta^n, j\in S_i} \lambda_{i,j} \left(\sum_{s:b^s=i} x^s_{i,j} - R_{out(i,j)} + R_{in(i,j)} \right) \\ &= \max_{\{\mathbf{x}\}} \left(\sum_{s\in\Theta^f} U_s(x^s) - \sum_{i\in\Theta^n, j\in S_i} \lambda_{i,j} \sum_{s:b^s=i} x^s_{i,j} \right) + \max_{\{P\}} \sum_{i\in\Theta^n, j\in S_i} \lambda_{i,j} \left(R_{out(i,j)} - R_{in(i,j)} \right) \end{aligned} \tag{6}$$

Solutions

- The congestion control problem

$$\max_{\{\mathbf{x}\in Co(\mathcal{R}),\mathbf{h}\leq 0\}} \left(\sum_{s\in\Theta^f} U_s(x^s) - \sum_{i\in\Theta^n, j\in S_i} \lambda_{i,j} \sum_{s:b^s=i} x^s_{i,j} \right) \tag{7}$$

- The scheduling problem

$$\max_{\{P\in\mathcal{P}\}} \sum_{i\in\Theta^n, j\in S_i} \lambda_{i,j} \left(R_{out(i,j)} - R_{in(i,j)} \right) \tag{8}$$

- The system governing law

$$\dot{\lambda}_{i,j}(t) = \left(\sum_{s:b^s=i} x^s_{i,j}(t) - R_{out(i,j)}(t) + R_{in(i,j)}(t) \right)^+_{\lambda_{i,j}(t)} \tag{11}$$

$$\dot{x^s_{i,j}}(t) = \left. \frac{\partial U_s(x^s)}{\partial x^s_{i,j}} \right|_{x^s(t)} - \lambda_{i,j}(t) + \delta(t) u\left(x^s_{i,j}(t)\right) + \sum_{m\in H^s} v_m(t) \left. \frac{\partial h_m(x^s)}{\partial x^s_{i,j}} \right|_{x^s(t)} \tag{13}$$

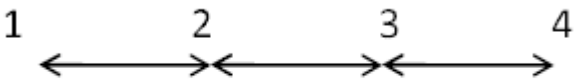
$$\hat{P}(t) = \arg\max_{P\in\mathcal{P}} \sum_{i\in\Theta^n, j\in S_i} \lambda_{i,j}(t) \left(R_{out(i,j)}(t) - R_{in(i,j)}(t) \right) \tag{16}$$

- We proved that (11), (13) and (16) converge to the optimal solutions if utility functions are concave

Simulation and Results

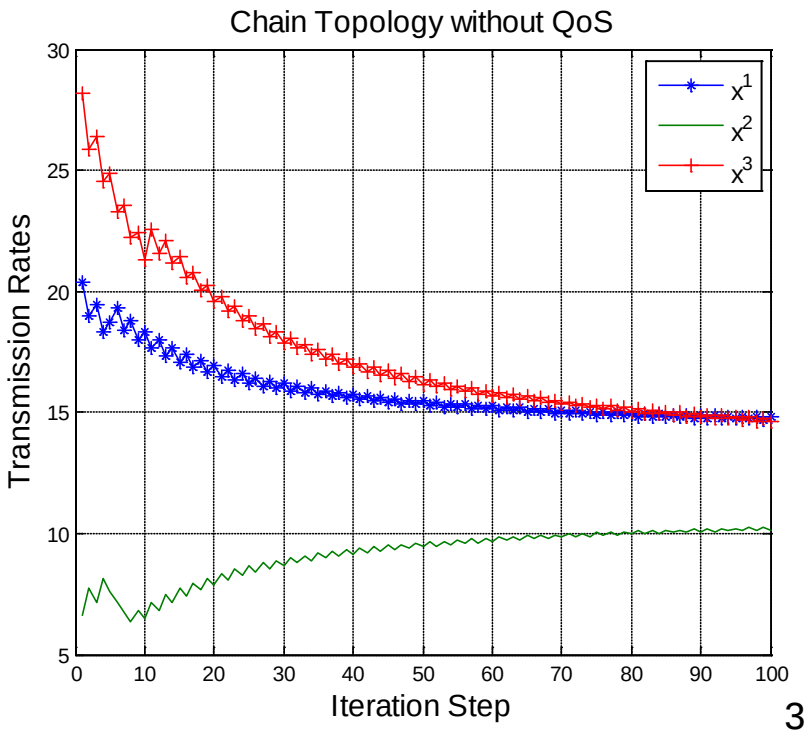
Chain topology:

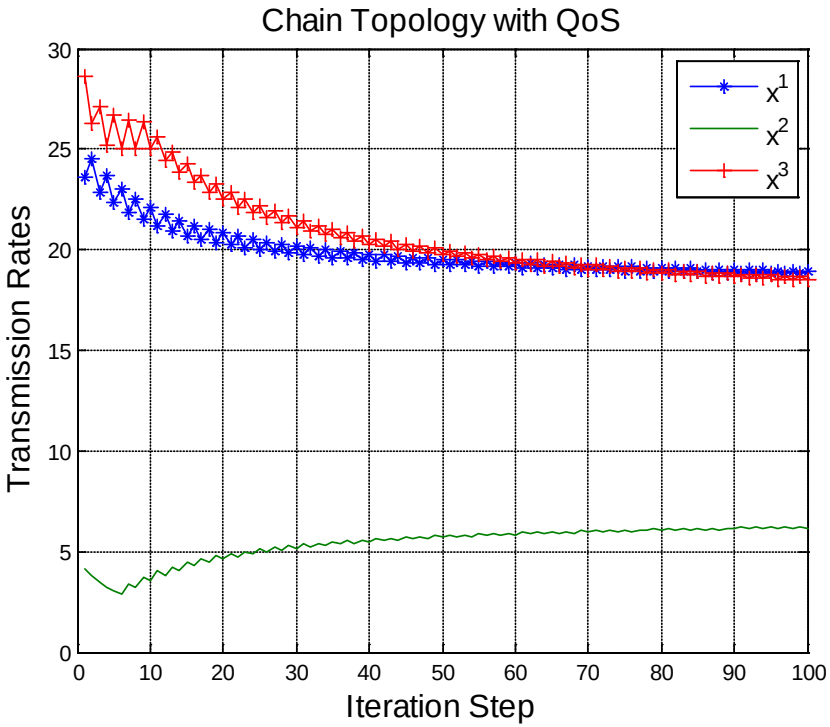
- $l_{2,3}$ is a bottleneck in the network expected to be smaller;
- The QoS constraint on x^1 can be satisfied



(a) Chain topology

Chain
Flow 1: $\{l_{1,2},\ l_{2,3}\}$
Flow 2: $\{l_{2,3},\ l_{3,4}\}$
Flow 3: $\{l_{3,4}\}$
QoS: $x^1 \geq 18$
v_m=1/3

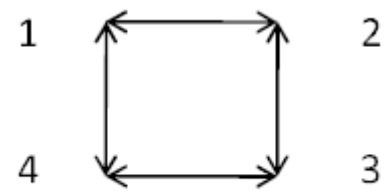




Results

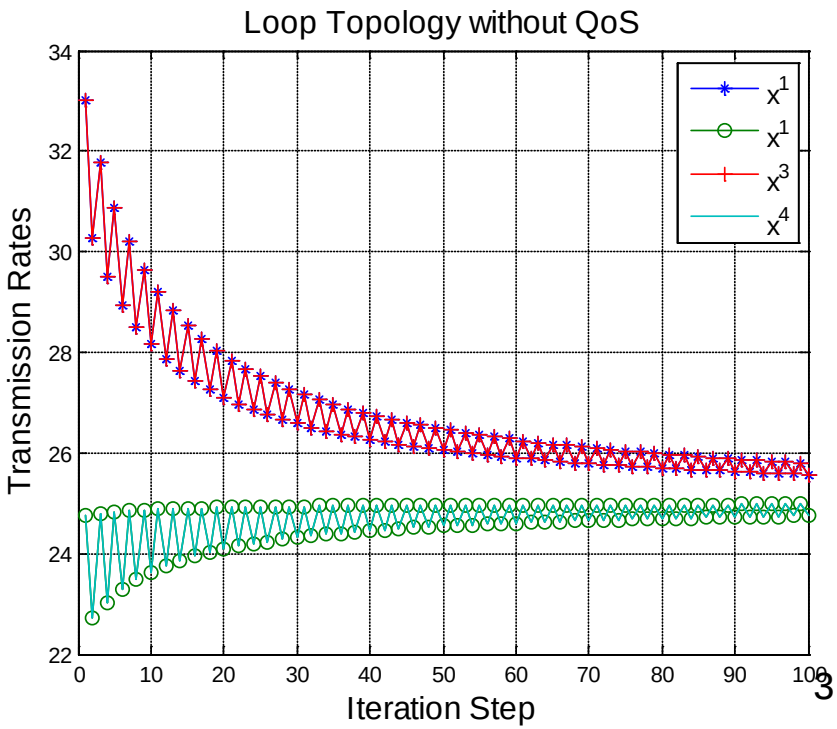
Loop topology:

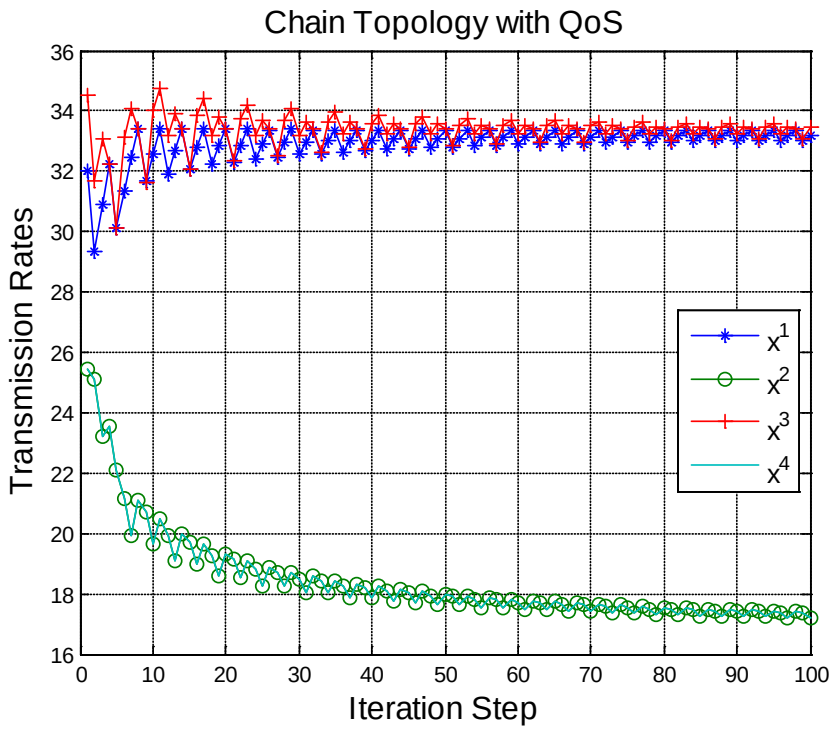
- The traffic is symmetric, so the transmission rates are also symmetric;
- The QoS constraint on x^1 can be satisfied



(b) Loop topology

Loop
Flow 1: $\{l_{1,2}\}$
Flow 2: $\{l_{2,3}\}$
Flow 3: $\{l_{3,4}\}$
Flow 4: $\{l_{4,1}\}$
QoS: $x^1 \geq 33$
v_m=1/2





Conclusions

Conclusions:

- Decomposition principle simplifies the NUM problem
- The per-next-hop queuing mechanism can reduce queuing overhead, compared to the per-destination queuing
- Network utility converges to the optimal
- The QoS constraints on transmission rates can be satisfied by sliding mode control

Thank you!