Reinforcement Learning Based Approach to Routing in Sensor Networks

CS 252 Project

Ambuj Tewari

ambuj@eecs.berkeley.edu

EECS, UC Berkeley

Problem

- Multihop routing in sensor networks.
- Many nodes, few sinks, dynamically changing asymmetric probabilistic links.
- Need a routing strategy which:
  - achieves good global routing performance.
  - involves local computation.
  - is adaptive to changes over time.
  - involves limited communication overhead.

Proposed Solution

- Treat problem as multi-agent partially observed RL problem.
- State of the network decomposes into states at nodes:
  - originating data queue
  - forwarding data queue
  - neighborhood table
- Actions: drop packet or select node to forward it to.
- Reward: local (based on link reliability) + global (broadcast by sink nodes)
- Decompose value function (& parameters) and use one of the standard learning techniques (Q-learning, Temporal Difference, Policy gradient)

Evaluation

- Will use Alec Woo's packet level simulator and TOSSIM.
- Evaluation criteria:
  - Is average global reward maximized?
  - What happens when nodes die? (count-to-infinity)
  - Do cycles form?
  - Time to convergence (if parameters are learned).
  - Topology stability (for non-randomized policies).

Expected Results

- Does multi-agent RL work for this problem?
- What is a good initialization strategy?
- Trade-off between exploration/exploitation.
- Relative performance of different reward functions and learning strategies.

Related Work

See annotated bibliography at:
`http://www.cs.berkeley.edu/~ambuj/cs252/`