Reinforcement Learning Based Approach to Routing in Sensor Networks

CS 252 Project

Ambuj Tewari

ambuj@eecs.berkeley.edu

EECS, UC Berkeley

Multihop Routing in Sensor Networks

- Routing in sensor networks has to tackle issues that don't usually arise in conventional communication networks.
  - Asymmetric probabilistic links
  - Resource (power, computation, memory) constrained nodes
  - Dynamic character of the network
- Approach to the problem taken in Alec Woo's thesis:

Routing strategy

$\Uparrow$ Logical connectivity graph

Neighborhood management

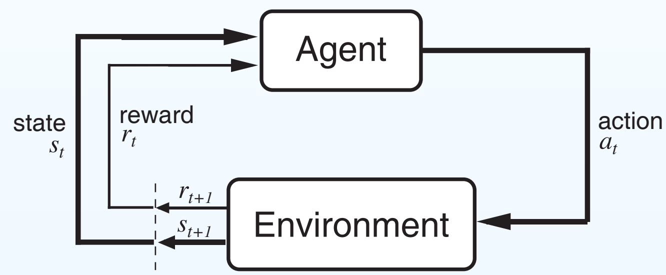
$\Uparrow$ Derived connectivity graph

Link Estimation

$\Uparrow$ Physical connectivity graph

Reinforcement Learning (RL)

- Basic theme of RL is learning to behave optimally under uncertainty based on feedback from the environment.



- Q-learning [Watkins 1989] is a well known method for learning a policy (mapping from states to actions).

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha[r_{t+1} + \max_a Q(s_{t+1}, a)]$$

- We intend to use this idea at the logical connectivity graph level. For $y \in \mathcal{N}(x)$,

$$Q(x, y) \leftarrow (1-\alpha)Q(x, y) + \alpha[\text{«reward»} + \max_{z \in \mathcal{N}(y)} Q(y, z)]$$

Reinforcement Learning

- The max Q-value of neighbor is obtained through periodic routing messages.
- The update requires very simple computation.
- Use ϵ-greedy strategy to route packets.
- Some relationship to Bellman-Ford type algorithms (set α = 1, reward = -1).
- Types of rewards:
  - -1 (hop count)
  - $1 - \frac{1}{p(x \rightarrow y)}$ (expected no. of retranmissions)
  - $\log p(x \rightarrow y)$ (likelihood of reaching destination)
- Count-to-infinity and cycle detection issues remain.

Evaluation

- Easy to simulate at the logical connectivity graph level (get idea of what obviously doesn't work).
- However, routing strategy has affect on levels below - such a simulation completely misses this.
- Use a packet-level simulator (Prowler?). Alec Woo also wrote his own (couldn't find it on the web).
- Most reliable is to test with real motes in a real environment.

Related Work

- Justin Boyan and Michael L. Littman: *Packet Routing in Dynamically Changing Networks: A Reinforcement Learning Approach*. In Advances in Neural Information Processing Systems 6, 1994

  Seed of idea. No convergence analysis of any kind. Toy simulations.

- Shailesh Kumar. *Confidence based Dual Reinforcement Q-Routing: an On-line Adaptive Network Routing Algorithm*. Master's Thesis, UT Austin, 1998

  Learning rate α is not constant but is a function of the confidence in the Q-value estimates. Confidence decays over time.

- Ying Zhang and Markus P. J. Fromherz. *Search-based adaptive routing strategies for sensor networks*. In AAAI-04 Workshop on Sensor Networks, 2004

  Asymmetric nature of links explicitly recognized. Theoretical analysis for the static case. Simulations using Prowler.