Poster Abstract: MDP Framework for Sensor Network Coordination

Shuping Liu[1] Anand Panangadan[2] Ashit Talukder[1,2,3] Cauligi S. Raghavendra[1]

[1] University of Southern California
Ming Hsieh Department of
Electrical Engineering
Los Angeles, CA 90089, USA
1-213-821-0871
{lius,raghu}@usc.edu

[2] Childrens Hospital Los Angeles
4650 Sunset Blvd.
Los Angeles, CA 90027, USA
1-323-361-2413
APanangadan@chla.usc.edu

[3] Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109, USA
1-818-354-1000
Ashit.Talukder@jpl.nasa.gov

1. INTRODUCTION

We consider a body area network application of monitoring a patient continuously [1,2]. In this application, several sensors are used in measuring physiological and metabolic readings of a patient. In such a system, the sampling rates have to be coordinated to maximize the life time of this sensor network system. We formulate the relationship between energy consumption, sensor sampling rates, and utility of coordination as a Markov Decision Process and compute a globally optimal policy for sensor sampling rates. We then present an entropy-based mechanism for communication between nodes to execute this global policy. We show preliminary results on simulated data to demonstrate that this distributed control framework is feasible for a limited number of sensors.

2. MDP FRAMEWORK

An MDP is defined as a 4-tuple (S, A, P, R). S is a finite set of states, in one of which the system exists. A is a set of actions that may be executed at any state. P is the probability of moving from one state to another after performing an action. R specifies the real-valued reward for performing an action in a state. A *policy* is defined as a function that determines an action for every state. If all the model parameters are known, the optimal policy can be computed by the Value Iteration algorithm.

In our application, we discretize all the relevant problem features. Actions are the sampling rates of the N sensors, $(a_1, a_2, \ldots, a_N)$. The *local state* of node N_i is represented by a state vector $s_i = (t_i, h_i, e_i)$. t indicates the number of control steps completed since the initial time. h is the application dependent measure of the criticality of the sensor readings (for instance, this could correspond to the health condition of a patient in a human health monitoring application). e is the amount of energy consumed. The *global state* is the joint local states of all the sensors, $S = (s_1, s_2, \ldots, s_N)$. The increase in control-step, change in event criticality, and fall in energy reserves are independent and hence we can define the probability of moving from global state s_i to s_j after taking action a

$$p(s_i, a, s_j) = p_T(t_i, t_j) p_H(h_i, h_j) \prod_{k=1}^{N} p_E(e_{k,i}, a_k, e_{k,j})$$

The component transition functions p_T,p_H, and $p_E$model the expected change in control-steps, event criticality, and energy. In particular, the rate at which energy is consumed by a sensor increases with the sampling rate. The reward function depends only on the sampling rates and the event criticality. It is proportional to the sampling rate and the event criticality. There is a large penalty if the system (i.e., *all* sensor nodes) runs out of power before some desired lifetime. (The MDP for a single sensor is described in our prior work [3].)

We compute the optimal sampling policy given the model parameters. If sensors can communicate their local state in every step, they can implement this optimal global policy during operation. As communication is expensive in terms of energy, we have formulated a method wherein every sensor maintains an estimate of the global state to utilize the global policy at every step. Sensors exchange the true local states only when it is determined that the value of information from the communication step (computed based on entropy of the global state estimate) exceeds a certain threshold. We express these concepts in terms of the above formulated MDP. Let S_{t+1}denote this set. S_{t+1} contains all the states that can be reached from a state in S_t by taking the action (sampling rates) specified by the policy π. (denoted by $\pi(s)$). Then, S_{t+1} is given by

$$S_{t+1} = \{s | \exists s' \in S_t, p(s', \pi(s'), s) > 0\}$$

The probability of the global state at $t + 1$ being a particular state $s \in S_{t+1}$ is given by

$$\Pr_{t+1}(s) = \sum_{s' \in S_t} \Pr_t(s') p(s', \pi(s'), s)$$

Supported in part by NSF grant no. 0615132

IPSN'09, April 13–16, 2009, San Francisco, California, USA.
ACM 978-1-60558-371-6

IPSN'09, April 13–16, 2009, San Francisco, California, USA.
Copyright 2009 ACM 978-1-60558-371-6/09/04...$5.00.

Note that the distribution of states at $t+1$ is dependent on the pre-computed policy, π. We use the information entropy of this distribution as the value of communication, V_t:

$$V_t = -\sum_{s \in S_{t+1}} \Pr_{t+1}(s) \log\left(\Pr_{t+1}(s)\right)$$

If V_t exceeds a pre-defined threshold (determined empirically in our current work), the sensor node triggers a communication step, which leads to exchange of local state information and exact knowledge of the global state. Otherwise, the local estimate of the global state is used.

3. SIMULATION RESULTS

We define the *system power outage* as the proportion of policy executions where *all* sensors ran out of power before the desired lifetime as a performance metric.

3.1 Communication

Figure 1 shows the system power outage with variation of the communication threshold for $N = 5$. The figure shows the trade-off between the energy cost of communication and the amount of available information.

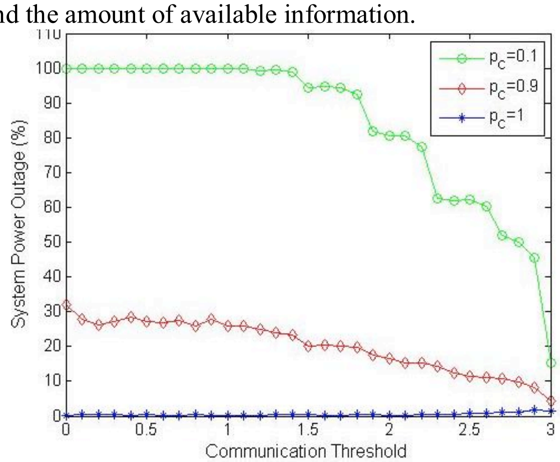

Figure 1: Effect of varying the threshold of communication on the system power outage. N=5.

3.2 Sensitivity to errors in model parameters

Figure 2 shows the effect of parameter error on the expected performance of the resulting policy. The results show that the performance of the policy gets better when the transition model is overestimated, and gets worse when it is underestimated.

3.3 Comparison with other policies

MDP policy is compared with other policies (Figure 3):

a. "Min": always sample at the lowest sampling rates.
b. "MDP-FC": MDP policy with free communication.
c. "MDP-CC": MDP policy with communication costs.
d. "Rand": Random policy.
e. "Heuristic": rate proportional to remaining energy.
f. "Max": always sample at the highest sampling rates.

Figure 3 shows that the MDP-FC policy results in a power outage probability as low as the minimum sampling rate case (the best possible) but while enabling the sensors to sample at higher sampling rates towards the end of the desired lifetime.

4. CONCLUSIONS

We have shown how the Markov Decision Process framework can be used for sensor coordination. This method is suitable for networks of relatively few sensors and where the computational capabilities and energy reserves at each node are limited.

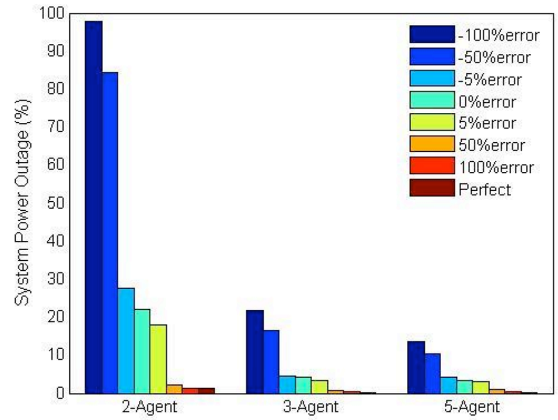

Figure 2: Probability of system power outage with varying amounts of model error ($\boldsymbol{N} = \mathbf{2, 3, 5}$).

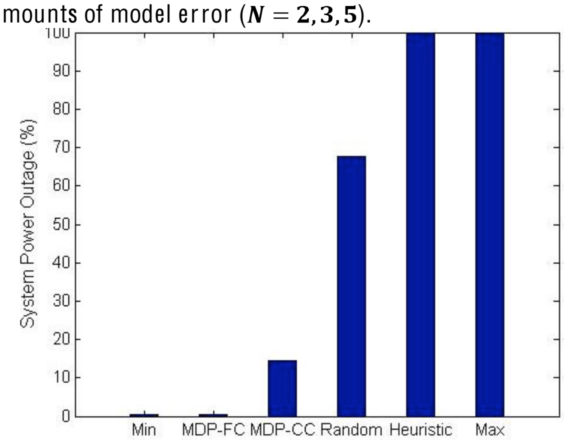

Figure 3: System power outage of MDP, Random, and Fixed policies ($\boldsymbol{N} = \mathbf{5}$).

5. REFERENCES

[1] C. S. Raghavendra, K. M. Sivalingam, and T. Znati, "Wireless Sensor Networks," Springer, 2006.

[2] M. Venugopal, K. E. Feuvrel, D. Mongin, S. Bambot, M. Faupel, A. Panangadan, A. Talukder, and R. Pidva, "Clinical Evaluation of a Novel Interstitial Fluid Sensor System for Remote Continuous Alcohol Monitoring," *IEEE Sensors Journal,* vol. 8, 2008.

[3] A. Panangadan, S. M. Ali, and A. Talukder, "Markov Decision Processes for Control of a Sensor Network-based Health Monitoring System," in *Proceedings of the Seventeenth Innovative Applications of Artificial Intelligence Conference*: AAAI Press, 2005.