# **EXPECTED-UTILITY-BASED SENSOR SELECTION FOR STATE ESTIMATION**

David Cohen, Douglas L. Jones \*

University of Illinois at Urbana-Champaign Department of Electrical and Computer Engineering Urbana, IL Sriram Narayanan

Kilby Labs Texas Instruments 13560 N Central Expressway Dallas, TX 75243

## ABSTRACT

Applications such as long-term environmental monitoring and large-scale surveillance demand reliable performance from sensor nodes while operating within strict energy constraints. There is often not enough power for sensors to make measurements all of the time. In these cases, one must decide when to run each sensor. To this end, we develop a one-step optimal sensor-scheduling algorithm based on expected-utility maximization. "Utility" is an applicationspecific measure of the benefit from a given sensor measurement. In sensing environments that can be modeled using a hidden Markov model, selecting the appropriate combination of sensors at each time instant enables maximization of the expected utility while operating within an energy budget. For some budgets, the utility-based algorithm shows more than 300% utility gains over a constant duty-cycle scheme designed to consume the same amount of energy. These benefits are dependent on the energy budget.

*Index Terms*— sensor management, utility maximization, energy management

### 1. INTRODUCTION

In state-estimation problems, sensor networks are often used to make state-observations. In sensor networks with strict energy constraints, there may not be enough power to run every sensor all of the time. There has been a fair amount of work on developing optimal or near-optimal sensor-scheduling algorithms. Often these sensor-selection problems are broken into two classes: geographically-based selection on distributed networks [1] and information-based selection [2]. We bypass this distinction by instead treating each sensor as having unique, state-dependent performance characteristics. Thus, the problem becomes that of selecting the k of n statedependent sensors that optimize some objective function. Because Hidden Markov Models (HMMs) have been shown to be simple yet accurate models for a variety of applications, we use them in our formulation as the state-evolution model.

Several different objective functions for sensor-selection have been suggested in the literature, including Chernoff and Kullback-Leibler distances [2], information gain [3], [4], and estimation error [5]. While engineering metrics such as these are the conventional optimization objectives, there is a large class of problems for which these metrics do not adequately express the true system-performance objective. For this reason, we propose using application-specific utility functions to provide the flexibility required for such problems. Unlike the stationary utility functions in [1] and [6], the utility in our formulation is both time- and state-dependent. Because the utility gained from sensing cannot be known *before* the sensor has run, we develop theory for sensor-selection based on maximizing *expected* utility.

In some applications, the cost to transmit information may be greater than the cost to run a sensor. By modeling data transmission as an activity with fixed cost and timedependent utility, our framework may additionally be used to schedule transmission times.

## 2. PROBLEM FORMULATION

We are interested in estimating the state,  $\theta_t \in \{0, 1, \ldots, m\}$ , of a Hidden Markov Model (HMM) with known parameters. We have *n* sensors from which we can collect noisy observations. At a given time, we can run one of  $2^n$  possible combinations of sensors. Because each combination may have unique performance characteristics, we treat them as individual units. We wish to pick the best "sensor collection" at each time. At time *t*, sensor collection  $k_t$  can make an observation  $y_t \in \mathbb{O}_k$ . Note that the observation space for each collection may be different. Every sensor collection has an energy cost,  $c(k_t)$ , associated with it. We assume that the observation model,  $P(y_t|k_t, \theta_t)$ , is known and independent accross time for every collection and every state. Because the true state is hidden, we use all past observations and sensor choices at each time to perform state estimation and sensor selection.

<sup>\*</sup>The authors acknowledge the support of the Multiscale Systems Center, one of six research centers funded under the Focus Center Research Program, a Semiconductor Research Corporation program.

That is, the information we have available at t is

$$I_t = (I_{t-1}, k_t, y_t), \quad I_{-1} = y_{-1},$$

Energy contraints restrict us from sensing with every sensor all the time. We wish to balance the utility of a sensor's measurement with its energy cost. Utility is a numerical or ordered metric that captures system performance in terms of value to the user. Bayes risk is a natural utility function in HMMs because it utilizes priors and allows for asymmetric weighting of performance errors. Under the Bayes risk formulation, utility is written as

$$U_t(\mathbf{I}_t) = \max_{i \in \{0,1,\dots,m\}} \sum_{\theta_t=0}^m R_{i\theta_t} P(\theta_t | \mathbf{I}_t).$$
(1)

 $R_{i\theta t}$  is the reward for deciding state *i* while the true state is  $\theta_t$ , and  $P(\theta_t | \mathbf{I}_t)$  is the posterior.

Because of the Markovian dynamics, the posterior can be updated at each time using only the previous posterior, current sensor choice, and current observation. This property eliminates the need for explicit knowledge of all previous observations and sensor choices [7]. That is, we have  $U_t(\mathbf{I}_t) = \overline{U}_t(P(\theta_t | \mathbf{I}_t), k_t, y_t)$ . The posterior at time t is calculated using Bayes rule in (2), and then propagated through the model for time t + 1 using (3).

$$P(\theta_t | \mathbf{I}_t) = \frac{P(y_t | k_t, \theta_t) \cdot P(\theta_t | \mathbf{I}_{t-1})}{P(y_t | k_t)}$$
(2)

$$P(\theta_{t+1} = j | \mathbf{I}_t) = \sum_{i=0}^m a_{ij} \cdot P(\theta_t = i | \mathbf{I}_t),$$
(3)

where  $a_{ij}$  is the one-step probability of transitioning from state *i* to state *j*.  $P(\theta_0|\mathbf{I}_{-1}) = \pi_0^{\theta}$ , the prior probability of state  $\theta_0$ .

We wish to maximize the total sum of utilities over a time horizon, N, subject to an energy constraint, B.

### 3. UTILITY MAXIMIZATION

We wish to select the sensor whose observation will provide the most utility. But because  $U_t(I_t)$  cannot be calculated *before* we have selected a sensor and made an observation, we maximize over *expected* utility.

$$\mathbf{E}[U_t|\mathbf{I}_{t-1}, k_t] = \sum_{y_t \in \mathbb{O}_k} P(y_t|k_t) \cdot U_t(\mathbf{I}_t)$$

Note that we have assumed the set of observations is discrete. An analogous equation can be developed for sets of continuous observations. We wish to find an optimal sensorscheduling policy,

$$\pi^{\star} = (\mu_0^{\star}(P(\theta_0|\mathbf{I}_0)), \dots, \mu_{N-1}^{\star}(P(\theta_{N-1}|\mathbf{I}_{N-1}))),$$

where  $\mu_t^{\star}(P(\theta_t|\mathbf{I}_t)) = k_t^{\star} \in \{0, 1, \dots, 2^n\}$  for all t, that solves

$$\max_{\pi} \sum_{t=0}^{N-1} \mathbf{E}[U_t | \mathbf{I}_{t-1}, k_t]$$
s.t. 
$$\sum_{t=0}^{N-1} \mathbf{E}[c(k_t)] \le B.$$
(4)

As with most constrained optimization problems, a Lagrange multiplier can be introduced in order to transform (4) into an unconstrained optimization problem. However, because the posterior probability is a *continuous* variable, it is very difficult to find an optimal policy. Furthermore, dynamic programming approximations that find optimal policies for a quantized posterior become computationally intractable as the number of states grows. We propose a one-step approximation inspired by the solution in [8]. For our problem, this solution turns out to be an approximation of the dynamic programming algorithm that does not consider future utility gains. At each time step, we find the sensor combination that solves

$$\min_{k_t \in \{0,1,\dots,2^n\}} - \mathbb{E}[U_t | \mathbf{I}_{t-1}, k_t] + \lambda c(k_t).$$
(5)

 $\lambda$  can be thought of as the minimum expected-utility-tosensor-cost ratio required to justify spending energy at a given time [9]. We choose  $\lambda$  to come as close to the energy budget, *B*, as possible. In almost every case we have tested, a  $\lambda^*$  can be found such that the desired energy budget is met within a small tolerance. This  $\lambda^*$  can be found by searching over all possible values of  $\lambda$  on a set of training data. The bisection method described in [10] is an efficient way to conduct this search. Once  $\lambda^*$  is found, the general procedure for a given time is as follows: select the optimal sensor-collection by solving (5), make the chosen observation, update the posterior using (2), determine the best state-decision by solving (1), and propagate the posterior through the model using (3).

In the next sections, we will discuss the results of our algorithm in a demonstration application.

#### 4. EXPERIMENTAL SETUP

As a test-application, we chose a room-occupancy detector for lighting-control (that is, estimate the state of occupancy of a room and turn the lights on/off accordingly). Our framework lends itself to this application for several reasons. First, office-activity has been successfully modeled using an HMM previously ([11], [12], and the references therein). We start with a simple two-state model (room occupied or unoccupied), but this model could be expanded to accommodate additional states. The second benefit to this application is the asymmetric cost of estimation errors; turning the lights off

SYSTEM		SENSORS	MIC	PIR	MIC + PIR
<i>a</i> <sub>11</sub>	0.985	cost	65	850	915
<i>a</i> <sub>00</sub>	0.993	$P_D$	0.24	0.77	0.83
$R_{11}$	1.25	$P_F$	0.07	0.04	0.11
$R_{10}$	-0.75				
$R_{00}$	0.25				
$R_{01}$	-1.5				

Table 1. System Parameters

while someone is in the room is far worse than turning the lights *on* when no one is in the room.

We created a 45-minute experiment in an office setting. A passive infrared (PIR) sensor and a microphone (MIC) provided data by which to estimate the state of the room. The estimate was then used to control the room-lighting.

We formed PIR observations by implementing an energydetector on 1.02 s frames. For the MIC data, we implemented a weighted GLRT on the power spectral density of 64 ms frames. When running both PIR and MIC sensors, we chose H1 if either the PIR or MIC observed H1. TABLE I summarizes the system parameters for our experimental setup.

The Markov transition probabilities,  $a_{11}$  and  $a_{00}$ , were determined using a transition frequency-counting technique on the training data. The Bayesian rewards,  $R_{i\theta}$ , were chosen to reflect the relative importance of missed detections and false alarms described at the beginning of this section. The energy-cost chosen for each sensor is based on its power consumption in  $\mu$ W. Because our algorithm is only affected by the *relative* costs of the sensors, it is not necessary to convert to Joules. For now we ignore the costs of the detection algorithms, but these could easily be included in the sensor costs.

#### 5. EXPERIMENTAL RESULTS

Figure 1 shows performance results for different energy budgets. The biggest performance benefits are observed in heavily energy-starved systems. This is because there is only enough energy to run the accurate sensors very occasionally. If these slots are wasted at the wrong times, the measurements will provide very little utility.

Figure 2 shows the evolution over time of the posterior probability of state 1 in a system constrained to consume 8.7% of the energy required for continuous operation of both sensors. The four state transitions from minutes 33 to 39 are shown. Figure 2 (a) shows the behavior of our algorithm, while (b) shows the behavior of running each sensor the same number of times as (a) at a constant duty cycle. The thick dashed line is an indicator of when the room was actually occupied. The thin solid line is the posterior probability that the room is occupied, and the thick solid line shows the corresponding state-decision. We assume the system turns the



**Fig. 1**. Utility increase over constant duty cycle scheme of the locally optimal scheme for different energy budgets.

lights on whenever it decides state 1.

Sensing causes large jumps in the posterior. When not sensing, the Markov transition probabilities cause the posterior to evolve toward the steady-state value of 0.3. In Figure 2 (b), each sensor is run at a constant duty-cycle. Because it trusts the sensor measurements without regard to state-structure or utility, we observe several costly missed detections.

Figure 2 (a) shows that our algorithm results in a "smart duty-cycle". Specifically, the PIR is used with a fast dutycycle in state 0, while the MIC + PIR is used with a slow dutycycle in state 1. This behavior makes it likely that a transition into state 1 will be detected quickly. Similarly, sensing rarely in state 1 decreases the possibility of a missed detection, and thus turning the lights off on someone. This rare-sensing results in more false alarms, which is exactly the tradeoff the utility-based reward-structure favors.

As of now no theoretical bounds have been developed on the utility achievable by a system with arbitrary parameters. However, if we assume a perfect sensor is available (i.e.  $P_D = 1$  and  $P_F = 0$ ), then we can compare the performance of our algorithm to that of an oracle that senses only at the true state-transition times. This oracle represents the maximum total utility possible under these conditions. Our algorithm approaches this limit quickly as a function of energy, achieving 93% of the maximum utility at an energy budget of 20%.

#### 6. CONCLUSION

Utility functions describe system performance in terms of value to the user. Because user value is an applicationspecific metric, utility functions have more flexibility as



Fig. 2. Evolution of the posterior probability and corresponding decision under (a) our locally optimal scheme and (b) a constant-duty-cycle scheme.

objective functions than standard metrics such as estimation error. Even though the one-step algorithm ignores the expected future payoff, we have seen that our method obtains over 90% of the maximum achievable utility while using only 20% of the total energy budget. The algorithm outperforms a constant duty-cycle method by significant margins across all enery levels. In addition, we observed that our algorithm automatically produces the desired tradeoffs between missed detections and false alarms on a real data set. Our results demonstrate great promise for using utility as an objective function in future applications.

## 7. REFERENCES

- J. Byers and G. Nasser, "Utility-based decision-making in wireless sensor networks," in *Mobile and Ad Hoc Networking and Computing*. IEEE, 2000, pp. 143–144.
- [2] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *submitted for pub.*, 2010.
- [3] E. Ertin, J. Fisher, and L. Potter, "Maximum mutual information principle for dynamic sensor query problems," in *Information Processing in Sensor Networks*. Springer, 2003, pp. 558–558.
- [4] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration for tracking applications," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 61–72, 2002.
- [5] S. Joshi and S. Boyd, "Sensor selection via convex op-

timization," *Signal Processing, IEEE Trans. on*, vol. 57, no. 2, pp. 451–462, 2009.

- [6] F. Bian, D. Kempe, and R. Govindan, "Utility based sensor selection," in *Proc. of the 5th Int Conf on Information processing in sensor networks*. ACM, 2006, pp. 11–18.
- [7] D.P. Bertsekas, *Dynamic programming and optimal control*, Athena Scientific, 1996.
- [8] H. Everett III, "Generalized lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, pp. 399–417, 1963.
- [9] D.G. Sachs and D.L. Jones, "Stochastic resource allocation for energy-constrained systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 3, 2009.
- [10] B.S. Krongold, K. Ramchandran, and D.L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *Communications, IEEE Trans. on*, vol. 48, no. 1, pp. 23–27, 2000.
- [11] N. Oliver, A. Garg, and E. Horvitz, "Layered representations for learning and inferring office activity from multiple sensory channels," *Computer Vision and Image Understanding*, vol. 96, no. 2, pp. 163–180, 2004.
- [12] B. Dong and B. Andrews, "Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings," in *Proc. Int. IBPSA Conf*, 2009.