ENERGY-EFFICIENT DETECTION SYSTEM IN TIME-VARYING SIGNAL AND NOISE POWER

Long Le, David M. Jun, and Douglas L. Jones

University of Illinois at Urbana-Champaign
Department of Electrical and Computer Engineering

ABSTRACT

In many detection applications with battery powered or energy-harvesting sensors, energy constraints preclude the use of the optimal detector all the time. Optimal energy-performance trade-off is therefore needed in such situations. In many applications, the signal and noise power may vary greatly over time, and this can be exploited to constrain energy consumption while maintaining the best possible performance. A detector scheduling algorithm based on the signal and noise power information is developed in this work. The resulting algorithm is simple due to its threshold-test structure and can be easily implemented with almost no overhead. A detection system with two detectors using the proposed scheduling scheme is estimated to greatly reduce the energy consumption for a wildlife monitoring application.

***Index Terms*—** Detection, low-power system, time-varying, scheduling, wildlife monitoring

1. INTRODUCTION

In many detection applications, energy is often an important constraint for system designers. For example, a remote wildlife monitor is usually required to operate solely on battery power for months before it can be recharged. As a result, the optimal but often sophisticated detector might not be feasible, and a trade-off must be made in order to maintain decent performance over the deployment lifetime. In many applications, the signal and noise power may vary greatly over time. For the wildlife monitoring application, signal and noise power varies as the animal of interest moves around or as other environmental noises come and go. This time-variation in the signal and noise power can be exploited to greatly reduce the energy consumption with little loss in performance using the new approach presented here.

The interest in reducing the complexity of detector's structure, and hence energy consumption, can be traced back to [1–3]. The common theme in the previous approaches is the generic optimization that do not take into account existing structure in certain detection applications. Later work [4, 5] exploits rare-event structure in their detection applications to enable even more aggressive energy saving. In addition, many detection applications also have time-varying signal and noise power structure that can be exploited. Early work that heuristically exploits this can be found in [6, 7].

Our contribution in this work is to derive the optimal detector scheduling policy that exploits the time-varying nature of signal and noise power for an energy-constrained detection system. The algorithm is a simple threshold test on the relative Bayesian risk between detectors. Since the threshold-test structure is simple, it adds no overhead to the overall system.

The authors would like to thank the support of the Multiscale Systems Center, one of six research centers funded under the Focus Center Research Program, a Semiconductor Research Corporation program.

2. FORMULATION

This section presents the framework for the detector scheduling problem. The input information to the scheduler is the 2-tuple random process $X_n = (P_n, Q_n)$ with P_n being the signal power and Q_n being the noise power. The output decision $U \in \{1, 2\}$ is the detector to use, assuming for simplicity that there are only two available detectors in the system. The scheduler itself is then a (possibly randomized) policy μ that maps X_n to 0 if the first detector is used and 1 if the second detector is used. From detection theory [8], a detector can be modeled by the equation

$$T_U(Y) \underset{H_0}{\overset{H_1}{\gtrless}} \tau_U$$

where H_0, H_1 are the two standard hypotheses: noise, and signal plus noise, respectively and Y is the noisy observation vector from one of the two hypotheses. The function T_U maps the given observation to a test statistic which is then thresholded by τ_U to decide which hypothesis was true. The test statistic for each detector is fixed so that the energy cost to compute them can be quantified and denoted by $e(U)$. Hence, at a particular time instance n, given X_n, the scheduling policy μ, and the threshold τ_U, the system risk and energy consumption (EC) can be defined as

$$R^{sys}(X_n, \mu, \tau_U) \triangleq \mu(X_n)R(X_n, \tau_1) + (1 - \mu(X_n))R(X_n, \tau_2)$$
$$EC(X_n, \mu) \triangleq \mu(X_n)e(1) + (1 - \mu(X_n))e(2)$$

where $R(X_n, \tau_U)$ is the respective detector risk that depends on the threshold and the signal and noise power. The goal then is to find the scheduling policy μ and thresholds τ_U that minimize the average system risk subject to an average energy consumption. Therefore, the optimization problem to be solved is

$$\begin{aligned} \min_{\mu,\tau_U} \quad & \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} R^{sys}(X_n, \mu, \tau_U) \\ \text{s.t.} \quad & \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} EC(X_n, \mu) \leq \beta \end{aligned} \tag{1}$$

where β is the average energy constraint.

In this work it is assumed that the process X_n is wide-sense stationary (WSS) and ergodic. Therefore the risk and energy consumption of the system are also WSS and ergodic, as the policy μ can not affect the natural process X_n. This allows us to convert the time average in (1) to the ensemble average with respect to the joint long-term statistics $p(\cdot)$ of X.

$$\begin{aligned} \min_{\mu,\tau_U} \quad & \mathbb{E}[R^{sys}(X, \mu, \tau_U)] \\ \text{s.t.} \quad & \mathbb{E}[EC(X, \mu)] \leq \beta \end{aligned} \tag{2}$$

Expanding (2) yields

$$\begin{aligned} \min_{\mu,\tau_1,\tau_2} \quad & \int dx\, p(x)\Big\{\mu(x)[R(x,\tau_1) - R(x,\tau_2)] + R(x,\tau_2)\Big\} \\ s.t. \quad & \int dx\, p(x)\Big\{\mu(x)[e(1) - e(2)] + e(2)\Big\} \leq \beta \end{aligned} \tag{3}$$

Observe in (3) that the thresholds only appear inside the detector risk so that the minimization over thresholds can be moved inside. This decouples (3) into two subproblems: optimization over thresholds, and optimization over policy.

2.1. Detector threshold optimization

The first subproblem is given by

$$\min_{\tau_U} R(x, \tau_U), \quad U = 1, 2$$

with

$$\begin{aligned} R(x, \tau_U) =& \pi_1 \int dy\, f_1(y) I(T_U(y|x) < \tau_U) \\ &+ \pi_0 \int dy\, f_0(y) I(T_U(y|x) \geq \tau_U) \end{aligned}$$

where $\pi_1 = \Pr(H_1)$, $\pi_0 = \Pr(H_0)$, and $f_0(y)$ and $f_1(y)$ are the observation's densities under the two hypotheses H_0 and H_1, respectively. $I(\cdot)$ denotes the indicator function.

Except for the special case where ratio between observation densities $f_1(y)/f_0(y)$ is the same as the test statistic $T_U(y)$, the optimal threshold is not the simple Bayesian threshold $\tau_U^* = \pi_0/\pi_1$ [8]. In general, τ_U^* need to be determined empirically, especially if a reliable observation model is not available.

2.2. Scheduler policy optimization

The second subproblem is given by

$$\begin{aligned} \min_{\mu} \quad & \int dx\, p(x)\Big\{\mu(x)[R(x,\tau_1^*) - R(x,\tau_2^*)] + R(x,\tau_2^*)\Big\} \\ s.t. \quad & \int dx\, p(x)\Big\{\mu(x)[e(1) - e(2)] + e(2)\Big\} \leq \beta \end{aligned}$$

Even though the above problem is a scheduling problem, it shares exactly the same structure as the well-known detection problem in the Neymann-Pearson lemma [9]. Applying the same machinery yields

$$\mu^*(x) = \begin{cases} 0 & \text{if } M(x) > \lambda \\ 1 \quad \text{w.p.}\rho & \text{if } M(x) = \lambda \\ 1 & \text{if } M(x) < \lambda \end{cases} \tag{4}$$

where $M(x) = \frac{R(x,\tau_1^*) - R(x,\tau_2^*)}{[e(2) - e(1)]}$ denotes the scaled relative Bayesian risk between two detectors, and $\lambda \in [0, \infty)$, $\rho \in [0, 1]$ are artificial variables [10]. The exact value of ρ and λ are determined from the energy constraint.

The fact that the optimal policy given in (4) is a simple threshold test is significant, as it implies that the addition of the scheduling module to the system adds virtually no extra overhead.

2.3. Robust scheduling policy

So far it is assumed that the signal and noise power information is available during the operation of the scheduler. However, in practice, signal and noise power need to be estimated [cf. Sec. 3.2]. How will this affect the scheduling policy obtained in (4)? In fact, it can be shown that the optimal policy structure in (4) is robust to estimated signal and noise power. However, the policy threshold needs to be modified to ensure that energy constraint is not violated even in the worst case [cf. Appendix A].

3. ALGORITHMS

This section complements the above discussion by describing the algorithm used to find the optimal policy threshold and the algorithm to estimate signal and noise power.

3.1. Optimal policy threshold

Without loss of generality, assume that $e(2) > e(1)$. Substituting the optimal policy mapping μ^* into the energy constraint yields

$$\begin{aligned} \mathbb{E}[EC(X, \mu^*(\lambda))] = & \int_{-\infty}^{\lambda} dt\, m(t) e(1) + \int_{\lambda}^{\infty} dt\, m(t) e(2) \\ & + m(\lambda)\big[(1-\rho)e(1) + \rho e(2)\big] \end{aligned}$$

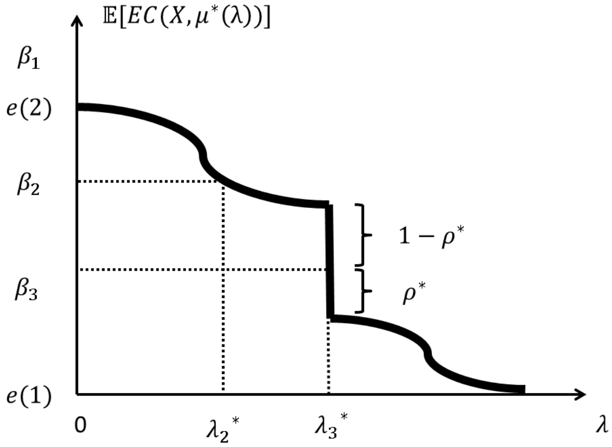

Fig. 1. Solution of λ^* and ρ

where $m(t) = \int dxp(x)I(M(x) = t)$.

In practice, $\mathbb{E}[EC(X, \mu^*(\lambda))]$ can be learned using training data. A sketch of $\mathbb{E}[EC(X, \mu^*(\lambda))]$ is given in Figure 1. From this, the solution for λ and ρ can be obtained. Depending on the energy constraint β:

- If $\beta > e(2)$ then the average energy constraint is redundant. This corresponds to the case $\beta = \beta_1$, hence $\lambda = \rho = 0$.
- If $e(1) \leq \beta \leq e(2)$ then the energy constraint can be satisifed with equality. If there is no point mass, i.e. $\beta = \beta_2$, then $\lambda = \lambda_2^*$ and $\rho = 0$. If there is point mass, i.e., $\beta = \beta_3$, then $\lambda = \lambda_3^*$ and $\rho = \rho^*$. The λ^* and ρ^* can be found using any root-finding method such as a bisection search.
- If $\beta < e(1)$ then the energy constraint is so stringent that it cannot be satisfied.

3.2. Signal and noise power estimation

Signal power P_n and noise power Q_n are estimated using recursive averaging. The desired update equations for noise power with smoothing coefficient α_Q are

$$\begin{aligned} H_0^n &: Q_{n+1} = \alpha_Q Q_n + (1-\alpha_Q)\|Y_n\|^2 \\ H_1^n &: Q_{n+1} = Q_n \end{aligned} \tag{5}$$

and for the signal power with smoothing coefficient α_P are

$$\begin{aligned} H_0^n &: P_{n+1} = P_n \\ H_1^n &: P_{n+1} = \alpha_P P_n + (1-\alpha_P)\Big[\|Y_n\|^2 - Q_n\Big] \end{aligned} \tag{6}$$

Notice that H_0^n and H_1^n are the hypotheses decided by the system and thus different from the true hypotheses H_0 and H_1. Let $\pi_1^n = \Pr(H_1^n)$. The update equation for π_1^n with smoothing coefficient α_π is

$$\pi_1^n = \alpha_\pi \pi_1^{n-1} + (1-\alpha_\pi)I(H_1^n)$$

Hence the two equations in (5) can be combined into

$$Q_{n+1} = \widetilde{\alpha}_Q^n Q_n + (1-\widetilde{\alpha}_Q^n)\|Y_n\|^2$$

where $\widetilde{\alpha}_Q^n = \alpha_Q + (1-\alpha_Q)\pi_1^n$. Similarly for (6)

$$P_{n+1} = \widetilde{\alpha}_P^n P_n + (1-\widetilde{\alpha}_P^n)\Big[\|Y_n\|^2 - Q_n\Big]$$

where $\widetilde{\alpha}_P^n = 1 - (1-\alpha_P)\pi_1^n$.

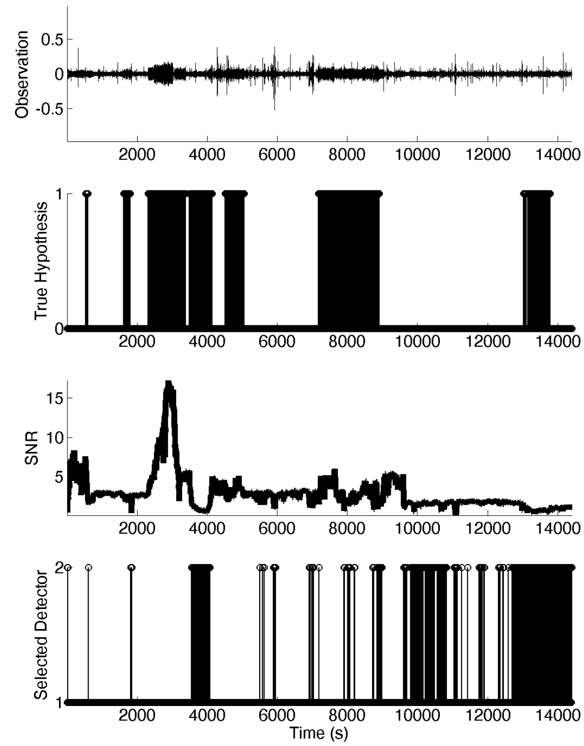

Fig. 2. Optimal detector scheduling on sample GCW data

4. APPLICATION: DETECTION OF GOLDEN-CHEEKED WARBLER

In this section we present the estimated performance of a detection system using the above scheduling algorithm. The target of detection are the calls of an endangered bird species named GCW [11]. The system employs two detectors. The first detector is the energy detector with simple implementation, namely $O(N)$ Multiply-ACcumulate (MAC) operations [12], where N is the size of the observation block Y.

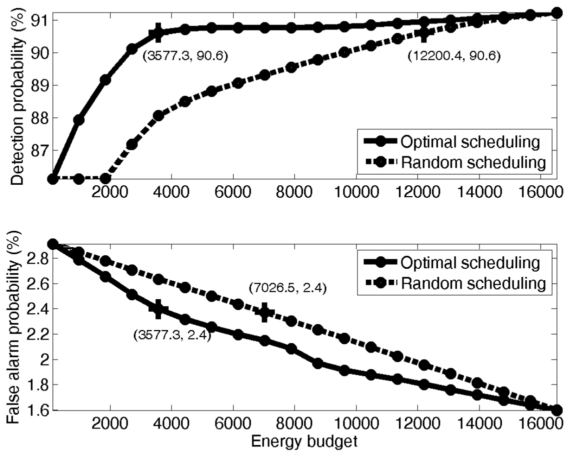

Fig. 3. Comparison between optimal scheduling and random scheduling over various energy budget.

The second detector is the *quadratic detector* [1]. Its implementation requires $O(N^2)$ MAC operations [12]. Therefore their respective cost $e(1)$ and $e(2)$ can be assigned to be N and N^2.

Figure 2 illustrates the system's operation on five hours of real GCW data, recorded by Professor Rama Ratnam from the Biology Department at the University of Texas at San Antonio [11]. The data is re-sampled at 16 kHz and processed in frames of size $N = 128$. It is then manually labeled as shown in the second window of Figure 2. The third window shows the tracked SNR using the algorithm in Section 3.2. The two detectors are managed by the scheduling algorithm described in Section 2.2. As can be seen from the fourth window, the second detector is run only when SNR is low while the first detector runs in the remaining time. The fraction of time between running the second detector and the first detector is determined by the energy budget. The operating point used in this case is labeled in Figure 3.

The optimal energy-performance curve is shown in Figure 3 as the solid line. The energy-performance curve of a random scheduling system, which is a straightforward but naive approach for this application, is shown in dash line. In this figure, the labeled points together illustrate the gap in energy budget required between the optimal scheduling system and the random scheduling system for the same level of desired performance. Namely, the optimal scheme is 3.5x more energy-efficient than the random scheme at the same level of detection probability. The number is 2x for false alarm probability. Furthermore, the optimal scheduling system's performance scales gracefully over an order of magnitude of the energy constraint.

5. CONCLUSION

In detection problems with an energy constraint, information about the time-varying signal and noise power can be used to optimally elect between the high performance but sophisticated detector and a simple but inexpensive one. The machinery to exploit this information by the scheduler is simple and hence no overhead cost is added. This results in a system that consumes much less energy while preserving adequate performance. An extension of this work is to implement the scheduling algorithm on an embedded detection hardware so that the energy-saving potential can truly be realized.

A. SOLUTION TO THE ROBUST FORMULATION

Assume that the estimated signal and noise power has a joint long-term statistic $g(\cdot)$ that is different from the true one $p(\cdot)$. For accuracy, $g(\cdot)$ needs to be modeled in (3) and the contaminated model [8] is proposed. For a given tolerance $\epsilon > 0$, this model assumes that the estimated signal and noise power are correct (i.e. drawn from the true density $p(\cdot)$) with probability $1 - \epsilon$ and incorrect (i.e. drawn from some bad density $h(\cdot)$) with probability ϵ [13]. Namely,

$$g(x) = (1 - \epsilon)p(x) + \epsilon h(x)$$

It is desirable to design a robust scheduling policy μ^R that can safeguard against the least favorable density g^L, i.e.

$$\begin{aligned} \min_{\mu} \max_{g} \quad & \int dx g(x) \Big\{ \mu(x)[R(x, \tau_1^*) - R(x, \tau_2^*)] + R(x, \tau_2^*) \Big\} \\ s.t. \quad & \int dx g(x) \Big\{ \mu(x)[e(1) - e(2)] + e(2) \Big\} \leq \beta \end{aligned} \tag{7}$$

The condition of von Neumann's minimax theorem can be verified easily to show that a saddle point solution (μ^R, g^L) to (7) exists. Observe that for any density g, the robust policy structure is still given by (4), i.e. $\mu^R = \mu^*$. In other words, the optimal policy in (4) is robust to estimated signal and noise power. The robust policy threshold λ^R, however, depends on the least favorable density that is yet to be found. Finding the least favorable g^L is the same as finding the least favorable h^L. This is equivalent to solving

$$\max_{h} \int dx h(x) \Big\{ \mu(x)[R(x, \tau_1^*) - R(x, \tau_2^*)] + R(x, \tau_2^*) \Big\} \tag{8}$$

Denote all the terms inside the curly bracket in (8) to be $f(x)$. Applying Schwarz's inequality to the integral in (8) reveals that the least favorable $h^L(x)$ is proportional to $f(x)$. Since h^L is a density, it is given by $h^L(x) = f(x)/\int f(t)dt$. Using this least favorable density, a robust policy threshold can be found using the approach in Section 3.1. The purpose of the robust policy threshold is to ensure that energy constraint is not violated even in the least favorable case.

B. REFERENCES

[1] H.V. Poor and C.I. Chang, "A reduced-complexity quadratic structure for the detection of stochastic signals," *The Journal of the Acoustical Society of America*, vol. 78, pp. 1652, 1985.

[2] L. Scharf and B. Van Veen, "Low rank detectors for gaussian random vectors," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 35, no. 11, pp. 1579–1582, 1987.

[3] Joseph M Winograd, S Hamid Nawab, and Alan V Oppenheim, "FFT-based incremental refinement of suboptimal detection," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*. IEEE, 1996, vol. 5, pp. 2479–2482.

[4] Paul Viola and Michael Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. IEEE, 2001, vol. 1, pp. I–511.

[5] David Jun, "An energy-aware framework for cascaded detection algorithms," M.S. thesis, University of Illinois at Urbana-Champaign, December 2010.

[6] S Gökhun Tanyer and Hamza Ozer, "Voice activity detection in nonstationary noise," *Speech and Audio Processing, IEEE Transactions on*, vol. 8, no. 4, pp. 478–482, 2000.

[7] K Srinivasan and Allen Gersho, "Voice activity detection for cellular networks," in *Speech Coding for Telecommunications, 1993. Proceedings., IEEE Workshop on*. IEEE, 1993, pp. 85–86.

[8] B.C. Levy, *Principles of signal detection and parameter estimation*, Springer Verlag, 2008.

[9] T.K. Moon and W.C. Stirling, *Mathematical methods and algorithms for signal processing*, vol. 1, Prentice hall New York, 2000.

[10] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, NY, USA, 2004.

[11] W.J. Leonard, J. Neal, and R. Ratnam, "Variation of type B song in the endangered golden-cheeked warbler (dendroica chrysoparia)," *The Wilson Journal of Ornithology*, vol. 122, no. 4, pp. 777–780, 2010.

[12] Y. Sung, L. Tong, and H.V. Poor, "Optimal and suboptimal detection of gaussian signals in noise: asymptotic relative efficiency," *Arxiv preprint cs/0507018*, 2005.

[13] P.J. Huber, "A robust version of the probability ratio test," *The Annals of Mathematical Statistics*, vol. 36, no. 6, pp. 1753–1758, 1965.