GROUP-ORDERED SPRT FOR DISTRIBUTED DETECTION

Yingwei Yao

University of Illinois at Chicago
Department of Electrical and Computer Engineering
851 S. Morgan St., Chicago, IL, 60607

ABSTRACT

We consider the problem of distributed detection in a large wireless sensor network. An adaptive data fusion scheme, group-ordered sequential probability ratio test (GO-SPRT), is proposed. This scheme groups sensors according to the informativeness of their data. Fusion center collects sensor data sequentially, starting from the most informative data and terminates the process when the target performance is reached. To analyze the average sample number, we establish the asymptotic equivalence between GO-SPRT, a multinomial experiment, and a normal experiment. Closed-form approximates are obtained. Our analysis and simulations show that, compared with fixed sample size test and traditional sequential probability ratio test (SPRT), the proposed scheme achieves significant savings in the cost of data fusion.

***Index Terms*—** Distributed detection

1. INTRODUCTION

We consider distributed detection in a wireless sensor network with one fusion center. Sensors take measurements, process the observations locally, and send the local summaries to the fusion center. The fusion center combines the local summaries received from sensors to reach a global decision. To reduce the communication cost of such distributed detection systems, two techniques have been explored in recent years: decentralized sequential detection [1, 2] and censoring [3–5].

In this paper, we propose a group-ordered sequential probability ratio test (GO-SPRT). Inspired by censoring schemes, we group sensors according to the informativeness of their data, with more informative data being collected first. Similar to sequential detection, the data collection process is terminated once the target performance is reached. The average sample number of GO-SPRT is analyzed. Our analysis and simulation show that this scheme achieves significant savings of communication cost over both fixed sample size tests and traditional distributed sequential detectors.

The rest of this paper is organized as follows. In Section 2, the signal model is described. The proposed scheme is presented in Section 3. In Section 4, we establish the asymptotic equivalence between a multinomial experiment and a multivariate normal experiment and obtain an approximate to the average sample number needed in GO-SPRT. In Section 5, we apply GO-SPRT to the problem of signal detection in Gaussian noise and present simulation results. Finally, we reach our conclusion in Section 6.

Notation: Upper and lower case bold symbols will be used to denote matrices and column vectors, respectively; $f(n) = O(g(n))$ means $\limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty$; if $0 < \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq \limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty$, then we say $f(n) = \Theta(g(n))$.

2. SIGNAL MODEL

We consider a distributed detection problem, where there are two hypotheses on the state of the environment:

$$H_0 : \boldsymbol{X}_n \sim P_0(\boldsymbol{X}_n), \qquad n = 1, \ldots, N, \tag{1}$$

$$H_1 : \boldsymbol{X}_n \sim P_1(\boldsymbol{X}_n), \qquad n = 1, \ldots, N, \tag{2}$$

where $\boldsymbol{X}_n \in \mathcal{X}$ contains the measurements taken by the nth sensor. The observations at different sensors are assumed to be independent given H_0 or H_1. With the communication constraint, the summary sent by a sensor must be selected from a finite set of messages. Suppose the summary sent by sensor n is $U_n = \gamma_n(\boldsymbol{X}_n)$, where γ_n is the decision rule adopted by sensor n. We know from detection theory that the optimal fusion center processing is to compute the log-likelihood ratio (LLR) of the received messages $\boldsymbol{U} = [U_1, \ldots, U_N]^T$:

$$\gamma_0(\boldsymbol{U}) = \sum_{n=1}^{N} \log \frac{P(U_n|H_1)}{P(U_n|H_0)}, \tag{3}$$

and compare it to an appropriate threshold. In this paper, we assume that all sensors use an identical local decision rule, which is asymptotically (for large N) optimal [6].

3. GROUP-ORDERED SPRT

We consider a wireless sensor network, where all sensors adopt an identical local decision rule $\gamma : \mathcal{X} \mapsto \{1, \ldots, 2K\}$.

Based on its observation $\boldsymbol{X}_n$, sensor n generates a summary $U_n = \gamma(\boldsymbol{X}_n)$. The distribution of U_n is described by the probability masses:

$$q_k(H_j) \triangleq P(U_n = k|H_j), \qquad k = 1, \ldots, 2K, \quad j = 0, 1. \tag{4}$$

We define $\Gamma_k \triangleq \{\boldsymbol{x} \in \mathcal{X} | \gamma(\boldsymbol{x}) = k\}$ as the set of observations associated with the summary message k. Denote $\mathcal{S}_k$ as the set of sensors, whose observations fall inside the region Γ_k, and $N_k \triangleq |\mathcal{S}_k|$ as the number of sensors in $\mathcal{S}_k$. In this paper, we assume that the following two conditions are satisfied:

Assumption 1: $q_k(H_0) \neq q_k(H_1)$ for some k;

Assumption 2: $q_k(H_j) > 0$ for all k and j.

Assumption 1 ensures that the summary messages contain some information about the state of the nature H and is satisfied for all but trivial local decision rules. Assumption 2 is also generally satisfied, except in the case of singular detection problems.

In a classical decentralized detection system, the fusion center collects summary messages from all sensors. To reduce the large communication and energy costs needed to collect all sensor summaries, we propose an adaptive fusion scheme: Group-Ordered Sequential Probability Ratio Test (GO-SPRT). In this scheme, the data collection process is divided into multiple stages. Sensors belonging to $\mathcal{S}_1$ and $\mathcal{S}_2$ send their summaries to the fusion center in the first stage, followed by the sensors in $\mathcal{S}_3$ and $\mathcal{S}_4$ in the second stage, and so on. At the end of each stage, the fusion center computes the global LLR based on the collected data. At the end of the k-th stage, the fusion center knows $N_1, N_2, \ldots, N_{2k}$ and can compute the following global LLR

$$\gamma_0^{(k)}(\boldsymbol{N}) = \sum_{i=1}^{2k} N_i \log \frac{q_i(H_1)}{q_i(H_0)} + \sum_{i=2k+1}^{2K} N_i \log \frac{\sum_{i=2k+1}^{2K} q_i(H_1)}{\sum_{i=2k+1}^{2K} q_i(H_0)}, \tag{5}$$

where we have denoted $\boldsymbol{N} \triangleq [N_1, \ldots, N_{2K}]^T$. The fusion center compares $\gamma_0^{(k)}(\boldsymbol{N})$ with thresholds $a < 0 < b$. If $\gamma_0^{(k)}(\boldsymbol{N}) \leq a$, the fusion center stops collecting data and chooses H_0; if $\gamma_0^{(k)}(\boldsymbol{N}) \geq b$, the fusion center chooses H_1; otherwise, the fusion center will start the data collection of the next stage.

GO-SPRT reduces the communication cost in two ways: i) it reduces the number of sensor transmissions, as demonstrated in this paper; ii) at stage k, each sensor that transmits needs only to send one bit indicating whether it belongs to $\mathcal{S}_{2k-1}$ or $\mathcal{S}_{2k}$, instead of the $\lceil \log_2(2K) \rceil$ bits needed in the traditional scheme.

4. AVERAGE SAMPLE NUMBER

To investigate the communication cost of GO-SPRT, we analyze the average sample number (ASN) needed under this scheme. Conditioned on H_j, the average sample number can be written as

$$ASN_j = \sum_{k=1}^{K} E\left[(N_{2k-1} + N_{2k}) \prod_{i=0}^{k-1} 1_{\{a < \gamma_0^{(i)}(\boldsymbol{N}) < b\}} \middle| H_j\right], \tag{6}$$

where 1_A is the indicator function of the subset A. Conditioned on H_j, $\boldsymbol{N} \sim \mathcal{M}(N; [q_1(H_j), \ldots, q_{2K}(H_j)])$ is a multinomial random vector. Direct evaluation of (6) is computationally costly when N and K are large. To analyze the behavior of ASN_j, we rely on the asymptotic equivalence between a multinomial experiment and a normal experiment.

4.1. Asymptotic Equivalence with a Normal Experiment

Since the sensors in $\mathcal{S}_1$ and $\mathcal{S}_2$ always send their summaries, $ASN_j = \Theta(N)$, that is, the fraction of sensors that report to the fusion center, $\varepsilon_j \triangleq ASN_j/N$, will not go to zero as N grows. We will refer to ε_j as the *fusion efficiency*. The following proposition provides an accurate approximation for ε_j for large N. Its proof is lengthy and has been omitted because of the space limit.

Proposition 1:

$$\varepsilon_j = \frac{1}{N} E\left[\sum_{k=1}^{K} (Z_{2k-1} + Z_{2k}) \prod_{i=0}^{k-1} 1_{\{a < \gamma_0^{(i)}(\mathbf{Z}) < b\}}\right] + O(\frac{1}{\sqrt{N}}), \tag{7}$$

where $\mathbf{Z} = [Z_1, \ldots, Z_{2K}]^T$ *is a Gaussian random vector with the same mean and covariance as* $\mathbf{N}$.

4.2. Single Boundary Crossing Approximation

The expectation in (7) involves a complicated multiple integral. To obtain a closed-form approximate for ε_j, we need further simplification. Intuitively, if $\gamma_0^{(j)}(\boldsymbol{Z}) \notin (a, b)$ for some $j > 0$, then it is unlikely that $\gamma_0^{(i)}(\boldsymbol{Z}) \in (a, b)$ for some $i > j$. Namely, the probability that $\gamma_0^{(i)}(\boldsymbol{Z})$ will cross the boundaries more than once is low. Motivated by this intuition, we investigate the following approximate:

$$\tilde{\varepsilon}_j = \frac{1}{N} \sum_{k=1}^{2K} E_j\left[Z_k 1_{\{a < \gamma_0^{(\lfloor \frac{k-1}{2} \rfloor)}(\boldsymbol{Z}) < b\}}\right]. \tag{8}$$

The approximation error of this single boundary crossing approximation is given by the following proposition, the proof of which is again omitted.

Proposition 2:

$$|\tilde{\varepsilon}_j - \hat{\varepsilon}_j| = o\left(\frac{1}{N^c}\right) \tag{9}$$

for all $c > 0$, *where*

$$\hat{\varepsilon}_j = \frac{1}{N} E\left[\sum_{k=1}^{K} (Z_{2k-1} + Z_{2k}) \prod_{i=0}^{k-1} 1_{\{a < \gamma_0^{(i)}(\mathbf{Z}) < b\}}\right]. \tag{10}$$

4.3. Closed-Form Approximation

From Propositions 1 and 2, we can see that ε_j can be accurately approximated by (8), leading to a closed-form approximation. To evaluate $\tilde{\epsilon}_j$, we note that under H_j, $[Z_k, \gamma_0^{(\lfloor \frac{k-1}{2} \rfloor)}(\boldsymbol{Z})]^T$ is a Gaussian random vector. Its mean is

$$\boldsymbol{\mu}_{kj} = [Nq_k(H_j), (\boldsymbol{c}^{(\lfloor \frac{k-1}{2} \rfloor)})^T \boldsymbol{\mu}_j]^T \triangleq [\mu_{kj}(1), \mu_{kj}(2)]^T, \tag{11}$$

where $\boldsymbol{c}^{(i)} \triangleq [c_1^{(i)}, \ldots, c_{2K}^{(i)}]^T$ with

$$c_k^{(i)} = \begin{cases} \log \frac{q_k(H_1)}{q_k(H_0)} & 1 \le k \le 2i, \\ \log \frac{\sum_{l=2i+1}^{2K} q_l(H_1)}{\sum_{l=2i+1}^{2K} q_l(H_0)} & 2i+1 \le k \le 2K. \end{cases} \tag{12}$$

Its covariance matrix is

$$\boldsymbol{\Sigma}_{kj} = \begin{bmatrix} \sigma_{kj}^2(1) & \rho_{kj}\sigma_{kj}(1)\sigma_{kj}(2) \\ \rho_{kj}\sigma_{kj}(1)\sigma_{kj}(2) & \sigma_{kj}^2(2) \end{bmatrix}, \tag{13}$$

where $\sigma_{kj}^2(1) = Nq_k(H_j)(1 - q_k(H_j))$,
$\sigma_{kj}^2(2) = (\boldsymbol{c}^{(\lfloor \frac{k-1}{2} \rfloor)})^T \boldsymbol{\Sigma}_j \boldsymbol{c}^{(\lfloor \frac{k-1}{2} \rfloor)}$, and

$$\rho_{kj} = \frac{q_k(H_j)(\boldsymbol{c}^{(\lfloor \frac{k-1}{2} \rfloor)})^T[-\boldsymbol{\mu}_j + N(2 - q_k(H_j))\boldsymbol{e}_k]}{\sigma_{kj}(1)\sigma_{kj}(2)}, \tag{14}$$

with $\boldsymbol{e}_k$ denoting a unit vector with the k-th entry equal to one.

After some algebra, we find

$$\tilde{\varepsilon}_j = \bar{\varepsilon}_j + O\left(\frac{1}{\sqrt{N}}\right), \tag{15}$$

where

$$\bar{\varepsilon}_j = \sum_{k=1}^{2K} q_k(H_j)\left[1 - Q\left(\frac{b - \mu_{kj}(2)}{\sigma_{kj}(2)}\right) - Q\left(\frac{\mu_{kj}(2) - a}{\sigma_{kj}(2)}\right)\right]. \tag{16}$$

4.4. Design of Local Decision Rule

In the above discussions, we have made no assumption on the local decision rule γ we use to categorize the sensor observations and divide the sensors into groups. A natural design option is the likelihood ratio quantizer (LRQ), which is known to be optimal for a large class of decentralized detection problems. Furthermore, we order the quantization cells such that

$$\left|\log \frac{q_1(H_1)}{q_1(H_0)}\right| \ge \cdots \ge \left|\log \frac{q_{2K}(H_1)}{q_{2K}(H_0)}\right|, \tag{17}$$

which assigns higher priority to more informative data. We assume no information on the prior probabilities of H_1 and H_0, so the quantization thresholds should be chosen to minimize $\max\{\varepsilon_0, \varepsilon_1\}$. To reduce the computational cost in optimizing the quantization thresholds, we rely on the person-by-person optimization (PBPO) approach: in each iteration, we update one threshold by keeping the other thresholds fixed and finding the new threshold value that minimizes the cost function.

5. SIGNAL DETECTION IN GAUSSIAN NOISE

In this section, we focus on the problem of signal detection in Gaussian noise to illustrate the application of GO-SPRT and its gain in fusion efficiency. Consider the following binary hypothesis testing problem:

$$\begin{aligned} H_0 &: X_n = -\theta + W_n, \qquad n = 1, 2, \ldots, N, \\ H_1 &: X_n = \theta + W_n, \qquad n = 1, 2, \ldots, N, \end{aligned} \tag{18}$$

where $\theta > 0$ is a constant signal and $\{W_n\}_{n=1}^N$ is a sequence of *i.i.d.* additive Gaussian noise samples with zero mean and variance σ^2. The local LLR at sensor n is $L_n = 2\theta X_n/\sigma^2$. Quantizing L_n is therefore equivalent to quantizing X_n. We will first consider the following uniform quantizer:

$$\Gamma_k = \begin{cases} [(K-1)\tau, \infty) & k = 1, \\ (-\infty, (1-K)\tau) & k = 2, \\ [(K - \frac{k+1}{2})\tau, (K + 1 - \frac{k+1}{2})\tau) & k > 2 \text{ \& } k \text{ is odd}, \\ [(-K - 1 + \frac{k}{2})\tau, (-K + \frac{k}{2})\tau) & k > 2 \text{ \& } k \text{ is even}, \end{cases} \tag{19}$$

where τ is a parameter to be chosen.

We apply the GO-SPRT algorithm, where we have set $\theta = 1$, $\sigma^2 = 10$, $a = -b = -20$, $\tau = 20/K$ and $N = 2000$. In Fig. 1, we plot both the approximate average sample number obtained using (16) and the results from Monte-Carlo simulation. We observe that the analytical results fit the simulation outcomes very well. Increasing the number of quantization levels improves the efficiency of the GO-SPRT scheme. The improvement is most significant when the number of quantization levels increases from 16 to 24. The gain appears to diminish as the number of quantization levels exceeds 100.

From Fig. 1, we observe that to obtain good efficiency with uniform quantizer, the number of quantization cells $2K$ needs to be large, which requires partitioning the data fusion process into many stages and may lead to large overhead. The necessary number of quantization cells can be reduced by using a quantizer with PBPO thresholds. From Fig. 2, we can see that person-by-person optimization of the quantization thresholds leads to significant improvements in the fusion efficiency.

We now compare the efficiency of the GO-SPRT with the fixed sample size (FSS) test and the traditional SPRT. We assume perfect quantization (no quantization error) for these two tests. Suppose the target false alarm probability is P_{F0}

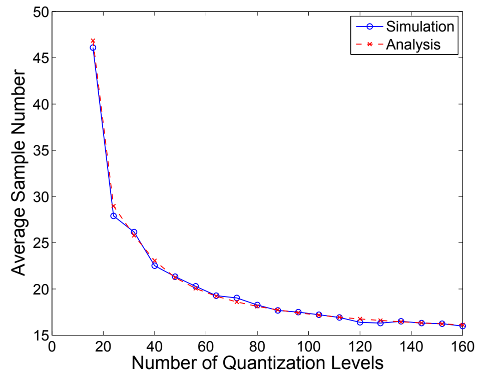

Fig. 1. Average Sample Number (under H_1) versus Number of Quantization Levels

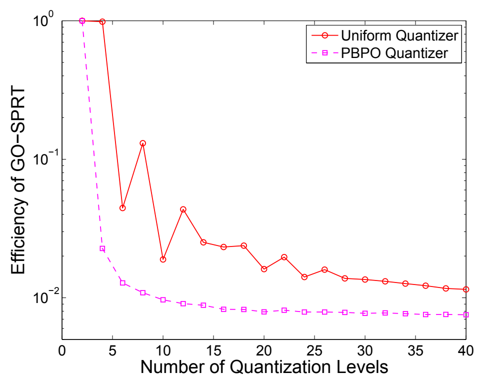

Fig. 2. Efficiency of GO-SPRT with Different Quantizers

and the target miss probability is P_{M0}. From Wald's approximation, we have $a \approx \log(P_{M0}/(1-P_{F0}))$ and $b \approx \log((1-P_{M0})/P_{F0})$. The number of samples needed under H_1 for the optimal FSS test is

$$N_{FSS} = \left\lceil \sigma^2[\Phi^{-1}(1-P_{F0}) - \Phi^{-1}(P_{M0})]^2/(4\theta^2) \right\rceil . \quad (20)$$

For $a = -b = -20$, we have $P_{F0} \approx P_{M0} \approx 2.0612 \times 10^{-9}$. Given $\theta = 1$ and $\sigma^2 = 10$, we have $N_{FSS} = 346$. For the traditional SPRT, simulation shows that the average sample number is approximately 102. On the other hand, simulation shows that GO-SPRT with PBPO quantizer requires only 45.4 samples for $K = 2$ and 16 samples for $K = 10$. Therefore, GO-SPRT with only a few quantization levels requires significantly fewer samples on average than both the FSS test and the traditional SPRT with perfect quantization.

Since N is limited, GO-SPRT may need to be truncated. However, under the above simulation setting, this effect is negligible. To see this, consider the worst case where $K = 1$ with sensor n sending $U_n = -1$ if $X_n \leq 0$ and $U_n = 1$ if $X_n > 0$. Given H_1, the truncation probability is

$$\begin{aligned} P_{T1} &\leq \sum_{m=0}^{\lceil \frac{N}{2} + \frac{b}{\log \frac{p}{q}} \rceil} \binom{N}{m} p^{N-m}(1-p)^m \\ &\leq \frac{q}{p-q}(2q)^N \left(\frac{p}{q}\right)^{\lceil \frac{N}{2} + \frac{b}{\log \frac{p}{q}} \rceil} \end{aligned} \quad (21)$$

where $q = Q\left(\frac{\theta}{\sigma}\right)$ and $p = 1-q$. For the above set of parameters, this bound yields $P_{T1} \leq 9.5 \times 10^{-24}$ for $N = 2000$. Compared with the target error probabilities, we can see that truncation has a minimal effect on GO-SPRT performance.

6. CONCLUSION

In this paper, we have developed a group-ordered sequential probability ratio test (GO-SPRT), which combines the advantages of sensor censoring and sequential probability ratio test. The fusion efficiency of this scheme is analyzed. Our analysis and simulation show that it has the potential to significantly reduce the communication cost of a distributed detection system. We have not addressed the design of transmission protocol within each fusion stage. This will be the subject of our future investigation.

7. REFERENCES

[1] V. V. Veeravalli, T. Basar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 433–442, 1993.

[2] Q. Cheng, P. K. Varshney, K. G. Mehrotra, and C. K. Mohan, "Bandwidth management in distributed sequential detection," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2954–2961, Aug. 2005.

[3] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Trans. Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[4] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Energy-efficient detection in sensor networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp. 693–702, Apr. 2005.

[5] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2668–2673, Nov. 2005.

[6] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing Vol.2: Signal Detection*, H. V. Poor and J. B. Thomas, Eds., pp. 297–344. JAI Press, Greenwich, CT, 1993.