10-1993

Some Comments on Distributed Detection and Estimation

R. Viswanathan
Southern Illinois University Carbondale, viswa@engr.siu.edu

Published in Viswanathan, R. (1993). Some comments on distributed detection and estimation. Conference Proceedings., International Conference onSystems, Man and Cybernetics, 1993. 'Systems Engineering in the Service of Humans', v. 3 669 - 671. doi: 10.1109/ICSMC.1993.385093

Some Comments on Distributed Detection and Estimation

R. Viswanathan
Department of Electrical Engineering
Southern Illinois University at Carbondale
Carbondale, IL 62901-6603
USA

Abstract **In this paper we discuss the question of information and accuracy attainable in distributed processing as compared to central processing. An example is presented where distributed detection suffers zero loss in performance as compared to central detection. In the same example, if the problem considered is one of estimation rather than detection, then it is shown that distributed estimation suffers a loss as compared to central estimation. This shows that the distributed detection and the distributed estimation problems cannot be considered on an equivalent footing. Some comments regarding an accuracy bound in detection problems are also provided.**

I. INTRODUCTION

In recent years, signal processing with distributed sensors is gaining importance. The relatively low cost of sensors, the inherent redundancy possible with multiple sensors, the availability of high speed communication networks and increased computational capability have spurred great research interest in this topic. Each sensor in a distributed sensor network (DSN) processes its observations and transmit only some condensed data to the fusion center. Therefore, it is expected that a distributed detection or estimation scheme suffers some loss in performance as compared to an optimal central scheme. In several situations, the loss happens to be small[1]. However, it is of interest to know what accuracy is attainable in a distributed scheme in relation to a central scheme. First an example is presented where distributed detection suffers zero loss in performance. In the same example, if the problem considered is one of estimation rather than detection, then it is shown that distributed estimation suffers a loss as compared to central estimation. This shows that the distributed detection and the distributed estimation problems cannot be considered on an equivalent footing. Next some comments regarding an accuracy bound in detection problems are provided.

II. A GENERAL RESULT IN DISTRIBUTED DETECTION

Let $Y_1, Y_2, .., Y_n$ denote the order statistics obtained from the observations $\{ X_1, X_2, .. X_n \}$, where X_i, i=1,2,..,n denotes the observation in the ith sensor of a DSN. Let us consider the situation where some inference concerning a real parameter θ is to be made. Consider the following hypotheses:

$$H_1: \ \theta \in \Omega_1 \quad \text{Vs.} \quad H_0: \ \theta \in \Omega_0 \qquad (1)$$

where Ω_1 and Ω_0 are some intervals on the real line. Point null hypotheses can be approximated by a vanishingly small interval. Consider the situation where for a given $f_{X_1, X_2, ..., X_n}(,, .., |\theta)$, the optimal (according to some criterion such as Baye's or Neyman-Pearson) central test is given by

$$\text{Decide } H_1 \text{ iff } Y_j > t \qquad (2)$$

where t is some threshold that satisfies the chosen criterion.

Lemma 1:

The distributed test equivalent to the central test in (2) is given as follows. Set

$$U_i = \begin{cases} 1 & if\ X_i > t \\ 0 & or\ else \end{cases}$$ and let the fusion center

$$\text{decide } H_1 \quad iff \quad \sum_{i=1}^{n} U_i \ \geq n - j + 1 \qquad (3)$$

The proof follows from the observation that the sets $\{Y_j > t\}$ and $\left\{ \sum_{i=1}^{n} U_i \ \geq n - j + 1 \right\}$ are one to one.

The optimal central test (2) implies that the order statistic Y_j is a sufficient statistic. In order to construct a test equivalent to the optimal central test, only the information whether Y_j exceeds t or not is required and not the sufficient statistic itself. As seen, the required information is obtainable from the $\{ U_i \}$. It should be noted that (2) and (3) are equivalent irrespective of whether the { X_i } are statistically independent or not. As an application of lemma 1, let us consider the following example.

2.1 An example

Consider the observations { X_i } to be iid uniform on (0, θ), $\theta > 0$, and consider the test:

$$H_0: \theta = \theta_0 \quad vs \quad H_1: \theta > \theta_0 \tag{4}$$

The largest order statistic is sufficient for θ and a uniformly most powerful size α test based on { X_i } is given by [2]:

$$\text{decide } H_1 \ \ iff \ \ Y_n > t \tag{5}$$

where t, chosen to achieve size α, is given by $t = \theta_0 \sqrt[n]{1-\alpha}$. According to lemma 1, the Boolean OR rule that declares hypothesis H_1 if at least one of the U_i's equals one, is also UMP.

III. DISTRIBUTED ESTIMATION

Consider the same model as in the above example but consider the problem to be the estimation of θ. Treating θ as an unknown constant, let us try to find an estimate of the parameter with some good properties such as unbiasedness and small variance[2]. Since Y_n is sufficient and complete, the uniformly minimum variance unbiased estimator of θ is given by

$$\theta_1^* = \frac{n+1}{n} Y_n \tag{6}$$

Hence (6) is a central estimator which is UMVUE. Simple calculation gives the variance of the estimator (6) as

$$Var(\theta_1^*) = \frac{\theta^2}{(n+2)n} \tag{7}$$

In the case of distributed estimation, a reasonable estimate of $\frac{1}{\theta}$ is more easily obtained than an estimate of θ. If an estimate of $\frac{1}{\theta}$ is to be obtained from the set of one bit quantized information { U_i }, some loss in performance as compared to a central estimator is certainly expected. Moreover, the performance loss depends on the choice of t in relation to the parameter θ which is unknown!. Let

$$T = \sum_i (1 - U_i) \tag{8}$$

$$\phi_1^* = \frac{T}{nt} \tag{9}$$

Above ϕ_1^* is an estimator of $\frac{1}{\theta}$. Straight forward but careful calculation shows the following:

$$Var(\phi_1^*) = \begin{cases} \dfrac{\theta - t}{nt\theta^2} & if\ t \le \theta \\ 0 & if\ t > \theta \end{cases} \tag{10}$$

$$Bias(\phi_1^*) = \begin{cases} 0 & if\ t \le \theta \\ \dfrac{t-\theta}{\theta t} & if\ t > \theta \end{cases} \tag{11}$$

The mean square error (MSE), which is the sum of the square of bias and variance , is given by

$$MSE = \begin{cases} \dfrac{(t-\theta)^2}{\theta^2 t^2} & t > \theta \\ \dfrac{(\theta - t)}{\theta^2 t\, n} & t \le \theta \end{cases} \tag{12}$$

A too conservative choice of t leads to one or the other kind of loss. For example if t is too close to 0, it is more likely that t is less than θ and therefore the variance of the estimator will be large, and if t is too large, there will be a penalty in terms of the bias. Of course, nothing better could be expected with such a coarse quantization of one bit. On the contrary, the same coarse quantization does not lead to any loss in the case of hypothesis testing, because of the existence of one to one mapping between (2) and (3). Another way to explain the difference in nature of the two problems (estimation and testing) is the following: in testing, we wonder whether θ is greater than θ_0 or not, whereas in estimation, we aim for the exact value of the parameter. Certainly in the former case, a coarser quantization may not be bad at all. This example also shows that in distributed estimation, a much finer quantization of sensor data would be required for better performance. Also, the one to one mapping between an optimal central test and a distributed test is rare. One example is lemma 1. In general there will be loss in performance in a distributed system as compared to a central system. A bound on this loss would determine the information and accuracy obtainable in a distributed system.

IV. INFORMATION AND ACCURACY ATTAINABLE IN DISTRIBUTED DETECTION?

Consider a binary hypothesis testing problem where a parameter θ under question belongs to two mutually exclusive intervals on the real line. Let the ith sensor observation be X_i, and let $\{ X_i \}$ be iid with a probability mass function (discrete case), or a continuous density function, denoted by $f(\ ;\theta)$. For the sake of convenience the continuous case is treated below. Identical results for the discrete case are obtained by replacing the integrals with summations. Let Z be a sufficient statistic and let $\psi(Z) \in [0,1]$ be the non randomized central test and let $u(U_i) \in [0,1]$, not a one to one mapping of $\psi(Z)$, be the distributed test based on the 0/1 decision variables U_i of the sensors. Then the probability of disagreement between ψ and u, P, is given by

$$P = \int |\psi - u| f_{UZ}(u,z)\,du\,dz = \int |\psi - u| f_{U|Z}(u|z) f_Z(z)\,du\,dz \tag{13}$$

The probability P can be considered as a measure of closeness of performances of a distributed scheme and a central scheme. It is observed that P is a function of the parameter θ. Since Z is sufficient, $f_{U|Z}(\ )$ is independent of θ. Assuming that the derivative of P with respect to θ exists and assuming that the regularity conditions in the Cramer-Rao lower bound on an unbiased estimator of θ are satisfied [3],

$$\frac{dP}{d\theta} = \int |\psi - u| f_{U|Z}(u|z) \frac{d f_Z(z)}{d\theta}\,du\,dz \tag{14}$$

Applying Cauchy-Schwartz inequality to (14), we obtain

$$\left(\frac{dP}{d\theta}\right)^2 \le \int |\psi - u|^2 f_{UZ}(u,z)\,du\,dz \cdot \int \frac{\left(\frac{d f_Z(z)}{d\theta}\right)^2}{f_Z(z)^2} f_{UZ}(u,z)\,du\,dz \le P\, I(f) \tag{15}$$

where I(f) is the Fisher's information given by

$$I(f) = E\left(\left(\frac{d}{d\theta}(\ln f_Z(z))\right)^2\right) \tag{16}$$

Because Z is sufficient, the Fisher's information of a sufficient statistic is same as the information in the whole sample[4]. $I(f)$ determines the sensitivity of a distribution to θ and (15) provides an inequality between P and its slope. Large changes in slope can happen only in regions where P is sufficiently large according to (15). The utility of (15) is somewhat limited because of the occurrence of P on either side of (15). Numerical evaluation of Pas a function of θ and that of the bound (15) can be done for specific examples. However, we have not carried out such an evaluation yet.

ACKNOWLEDGMENT

The author would like to acknowledge partial funding for this work from the Office of Naval Research, USA (for task s400026srb01).

REFERENCES

[1] J.N. Tsitsiklis, "Decentralized Detection," in Advances in Statistical Signal Processing, Vol. 2, Signal Detection, H.V. Poor and J.B. Thomas, Eds. Greenwich, CT: JAI press, 1990.
[2] E.L. Lehmann, *Testing Statistical Hypotheses,* John Wiley &Sons, New York, 1986.
[3] E.L. Lehmann, *Theory of Point Estimation,* John Wiley & Sons, New York, 1983.
[4] S. Zacks, *Parametric Statistical Inference,* Pergamon Press, Oxford, 1981.