An Estimator for Multi-Sensor Data Fusion

Chandrasekhara Thejaswi, Viswanath Ganapathy, Ranjeet K. Patro, Manik Raina, Subhas K. Ghosh
Honeywell Technology Solutions Laboratory
151/1, Doraisanipalya, Bannerghatta Road,
Bangalore, India, 560076
Email:Chandrashekhara.Thejaswi@honeywell.com

***Abstract*— In this paper, we examine binary hypothesis testing and parameter estimation problem in a sensor network. We address the problem of detection and also the estimation of the underlying parameter at the fusion center by optimally combining the test statistics sent by different sensors. We make no assumptions on the noise statistics at the sensor nodes except that their first and second order statistics are known. We show that the proposed method is optimal as an estimator as well as a detector.**

I. Introduction

In recent years, signal processing techniques for sensor networks has been gaining immense attention of the research community. Sensors being tiny, low cost devices find applications in many areas like surveillance, industrial monitoring, environmental monitoring etc.

The critical challenges in the operation of a sensor network lie in the efficient exchange of information, optimal collaboration among the nodes to infer and gather information about the physical world. As we know, these challenges can be accomplished by efficient statistical signal processing algorithms which are broadly classified into *detection* and *estimation*. Researchers have been diligently pursuing these areas to come out with efficient algorithms for sensor networks. There is a significant collection of literature in these fields. Distributed detection has drawn the attention of researchers toward it. Survey papers [2], [1] provide excellent references to the earlier work in this field. A comprehensive review of the theories for decentralized detection is given in [2]. It is also shown that, under the assumption of conditional independence of the sensors' observations, local decisions made by the sensors obey likelihood ratio test. If there is no assumption of conditional independence among sensor's observations, then the problem of finding optimal decision strategy is NP-complete. Decentralized detection problems under cost constraints are also addressed in several works [3] etc. The problem of estimation has also been studied by several people [4]. Paper [4] addresses the problem of optimal estimation of unknown parameter at the fusion center from the noisy estimates obtained by the sensors. The problem is addressed under the communication and bandwidth constraints.

We analyze the problem of both detection and estimation at the fusion center using the noisy data obtained from the different sensors. The novelty of our approach lies in the fact that both the detection and estimation performed simultaneously and optimally. We also make no assumption on the noise statistics of the sensors except that the fusion center is cognizant of just first and second moments of them. These kind of problems commonly find their application in heterogeneous sensor network or the places where there is a need to not only detect the event but also to estimate the parameter associated with the event. Typical examples of such scenario could be temperature monitoring, where it is required to keep regularly the track of its value and to rise an alarm when it hits certain threshold. This requires both estimation of the parameter and also the detection of an event say, temperature shoot-up.

The rest of the paper is organized as follows. In Section II, we state the system model and the assumptions made. In Section III, we discuss about the binary hypothesis testing problem with the fusion center performing yhe optimal detection using sensor observations corrupted by noise. Section IV deals with the optimal estimation of the parameter using sensor observation which turns out to be test statistic itself. Finally, we conclude the paper in section VI.

II. The Model

We consider a parallel configuration of sensors sampling an event with a fusion center to detect a binary phenomenon. We assume that the sensors do not communicate with each other and they communicate only with the fusion center. We are concerned about the binary hypothesis testing of the phenomenon and estimation of the underlying signal at the fusion center based on the sensor observations. The hypotheses H_0 and H_1 have P_0 and P_1 respectively as their priors. Let there be K sensors and $\mathbf{X}_k, k = 1, \ldots, K$ be the $N \times 1$ noisy observation vector available at each sensor.

Assuming the state of an event to be stationary in the observation period, we have for the sensor k,

$$\mathbf{X_k} = Z + \mathbf{v_k}, \qquad (1)$$

where $\mathbf{v}_k$ is the additive noise vector at k^{th} sensor. We assume that noise at each sensor is generic identically distributed, uncorrelated in time with zero mean and variance σ_{vk}^2. It is assumed that fusion center has the knowledge of the second moment of sensor noise. At each epoch, sensor i gathers N noisy samples/observations of the phenomenon and generates its estimate as

$$\widehat{X}_k = \frac{1}{N} \sum_{n=0}^{N-1} X_k(n).$$

The estimate $\widehat{X}_k$ of the sensors are unbiased. Therefore, for large N, we can approximate the estimate to be Gaussian distributed as $\widehat{X}_k \xrightarrow{d} \mathcal{N}(Z, \frac{\sigma_{vk}^2}{N})$, Z being the signal level corresponding to the underlying hypothesis. Each sensor transmits its own estimate to the fusion center. Based on this information, fusion center obtains an optimal estimate of the parameter and also arrives at the global decision in favor of the occurred phenomenon (H_0 or H_1).

III. Detection with noisy observations

In this section we will see how the sensor information is used by the fusion center as the test statistic for hypothesis testing. Fusion center views the data sent by each sensor $\widehat{X}_k$ as a observation of the parameter corrupted by the estimation noise and the noise at the fusion center. At the fusion center we have

$$\mathbf{Y} = \widehat{\mathbf{X}} + \mathbf{v}_0 \tag{2}$$

where $\mathbf{Y} = [Y_1, \ldots, Y_K]$ is the observation vector received by the fusion center, $\widehat{\mathbf{X}} = [\widehat{X}_1, \ldots, \widehat{X}_K]$ is the estimates transmitted by sensors. $\mathbf{v}_0$ is the noise at the fusion center assumed to be AWGN with variance $\sigma_0^2.\mathbf{I}_K$. From (1) and (2), assuming conditional independence, the distribution of Y given Z is

$$f(\mathbf{Y}|Z) = \frac{1}{(2\pi)^{\frac{K}{2}} . \prod_{k=1}^K \sigma_k^2} e^{-\sum_{k=1}^K \frac{(Y_k - Z)^2}{2\sigma_k^2}} \tag{3}$$

where $\sigma_k^2 = \frac{\sigma_{vk}^2}{N} + \sigma_0^2$. According to standard Neyman-Pearson(NP) criteria, the optimal testing rule that can maximize the probability of a detection constraining on false alarm is the likelihood ratio test (LRT) which says for $P_F = \alpha$

$$\text{D}ecide\ H_1\ if$$

$$T(\mathbf{Y}) = \frac{f(\mathbf{Y}|H_1)}{f(\mathbf{Y}|H_0)} > \gamma \tag{4}$$

where γ is chosen such that the false alarm

$$P_F = \int_{\mathbf{y}:T(\mathbf{y})>\gamma} f(\mathbf{y}|H_0)\, d\mathbf{y} = \alpha$$

Fusion center computes the likelihood ratio which is given by

$$T(\mathbf{Y}) = \frac{f(\mathbf{Y}|H_1)}{f(\mathbf{Y}|H_0)} = \frac{e^{-\sum_{k=1}^K \frac{(Y_k - \theta_1)^2}{2\sigma_k^2}}}{e^{-\sum_{k=1}^K \frac{(Y_k - \theta_0)^2}{2\sigma_k^2}}}$$

Taking logarithm and upon simplification, we have,

$$\ln(T(\mathbf{Y})) = (\theta_1 - \theta_0) \sum_{k=1}^K \frac{1}{\sigma_k^2} \left(Y_k - \frac{\theta_0 + \theta_1}{2} \right)$$

Now, the fusion center's strategy is to decide in favor of H_1 if, $\ln(T(\mathbf{Y})) > \gamma$. That is decide in favor of H_1, if

$$T'(\mathbf{Y}) = \frac{\sum_{k=1}^K \frac{1}{\sigma_k^2} Y_k}{\sum_{k=1}^K \frac{1}{\sigma_k^2}} > \gamma',$$

where $\gamma' = \frac{\ln \gamma}{(\theta_1 - \theta_0)} + \frac{\theta_0 + \theta_1}{2(\theta_1 - \theta_0)}$. Let us interpret the test statistic test statistic $T'(\mathbf{Y})$ which is a weighted sum of the observations. It says that the observations which correspond to a more noisy and hence unreliable, must be given less weight.

To determine the detection performance we note that the that the test statistic $T'(\mathbf{Y})$ is Gaussian under each hypothesis. Also,

$$\begin{aligned} E[T'(\mathbf{Y})|Z] &= \frac{\sum_{k=1}^K \frac{1}{\sigma_k^2} E Y_k}{\sum_{k=1}^K \frac{1}{\sigma_k^2}} \\ &= \frac{\sum_{k=1}^K \frac{1}{\sigma_k^2} E \widehat{X}_k}{\sum_{k=1}^K \frac{1}{\sigma_k^2}} \\ &= Z \end{aligned}$$

and

$$\begin{aligned} var[T'(\mathbf{Y})|Z] &= var\left(\frac{\sum_{k=1}^K \frac{1}{\sigma_k^2} Y_k}{\sum_{k=1}^K \frac{1}{\sigma_k^2}} \right) \\ &= \frac{\sum_{k=1}^K \left(\frac{1}{\sigma_k^2}\right)^2 var \widehat{X}_k}{\left(\sum_{k=1}^K \frac{1}{\sigma_k^2}\right)^2} \\ &= \left(\sum_{k=1}^K \frac{1}{\sigma_k^2} \right)^{-1} \end{aligned}$$

we have then [**?**],

$$P_F = Q\left(\frac{\gamma' - E[T'(\mathbf{Y})|\theta_0]}{var[T'(\mathbf{Y})|\theta_0]} \right) = Q\left((\gamma' - \theta_0) \sum_{k=1}^K \frac{1}{\sigma_k^2} \right),$$

and

$$P_D = Q\left(\frac{\gamma' - E[T'(\mathbf{Y})|\theta_1]}{var[T'(\mathbf{Y})|\theta_1]} \right) = Q\left((\gamma' - \theta_1) \sum_{k=1}^K \frac{1}{\sigma_k^2} \right).$$

IV. Estimation with noisy observations

Given the noisy observations from the sensors, fusion center is confronted with the problem of estimation of the event parameter. That is given $Y_1, \ldots, Y_K$, the objective of the fusion center is to estimate Z. Given its little knowledge about the parameter and the statistics of the sensor observations, the common approach for fusion center is to resort for the best linear unbiased estimator (BLUE) [**?**]. That is, estimate of Z

$$\widehat{Z} = \sum_{k=1}^K \alpha_k Y_k,$$

where $\alpha_k,\ k = 1, \ldots, K$ are chosen such that $E\widehat{Z} = Z$ and the variance is minimized.

It can be shown that in our case that $\alpha_k = \frac{\frac{1}{\sigma_k^2}}{\sum_{k=1}^K \frac{1}{\sigma_k^2}}$. Thus, the optimal linear estimate we have is

$$\widehat{Z} = \frac{\sum_{k=1}^K \frac{1}{\sigma_k^2} Y_k}{\sum_{k=1}^K \frac{1}{\sigma_k^2}}$$

which is again the test statistic $T'(\mathbf{Y})$. The estimation error is given by the variance of the statistic which is $Error(\widehat{Z}) = \left(\sum_{k=1}^{K} \frac{1}{\sigma_k^2}\right)^{-1}$.

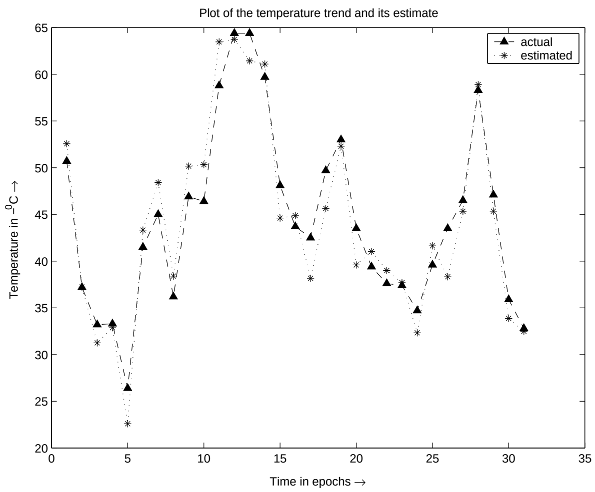

Fig. 1. Estimation of the Temperature Trend

V. Simulation Results

In this section we describe the simulation of our approach. We considered a typical temperature profile. Three sensors with different noise statistics observe the temperature trend which estimate the temperature finally report it to the fusion center. The system model was as follows $v(1) \sim Uniform(-.5, +.5)$, $v(2) \sim \mathcal{N}(0, .1), v(3) \sim Uniform(-.7, +.7)$. The number of observations per slot $N = 10$. The on board processing signal level corresponding to the temperature in the system is $0 - 0.1V$. Fig 1 shows the plot of actual and estimated temperature data. We can see that the even when the SNR (SNR was around 1-3dB) is low, the estimators capture the overall variation in the temperature trend and it shows that weighting indeed makes the estimator robust.

VI. Conclusions

In this paper, we discussed the problem of optimal estimation from the senosr data where each sensors are subjected to unidentical noise process. We also showed that the same estimate is a robust statistic for the detection.

References

[1] R.S. Blum, S.A. Kassam and H.V. Poor, “Distributed Detection with Multiple Sensors: Part II - Advanced Topics,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64-79, January, 1997.

[2] J.N. Tstsiklis, “Decentralized Detection,” *Advances in Statistical Signal Processing*, Vol 2, pp. 297-344, JAI Press, 1993.

[3] Jin-Jun Xiao, Zhi-Quan Luo, “Decentralized detection in a bandwidth constrained sensor network,” *IEEE GLOBECOM ’04*, vol. 1, pp. 123- 128, 2004.

[4] Zhi-Quan Luo, “Universal Decentralized Estimation in a bandwidth constrained sensor network,” *IEEE Transactions on Information Theory*,Vol. 51, No. 6, June 2005.