STATISTICAL RANKING OF SENSOR OBSERVATIONS FOR CENTRALIZED DETECTION WITH DISTRIBUTED SENSORS

Eric Ayeh , *Student Member, IEEE,* Kamesh Namuduri, and Xinrong Li, *Senior Members, IEEE*

***Abstract*—This paper investigates data reduction strategies from signal processing perspective for centralized detection with distributed sensors. We consider a deterministic source observed by a network of sensors and develop an analytical strategy for ranking sensor transmissions based on their test statistics. The benefit of the proposed strategy is that in certain scenarios, the decision to transmit or not to transmit to the fusion center can be made at the sensor level, resulting in significant savings in transmission costs. We derive a theoretical bound on the number of sensor transmissions saved. Our results complement existing results in the literature. We simulate the proposed strategy and demonstrate its benefits over the unconstrained energy approach.**

***Index Terms*—wireless sensor networks, distributed target detection, statistical ranking, sequential probability ratio test**

I. INTRODUCTION

IN a typical wireless sensor network (WSN) application such as target detection, several sensors are deployed in the field. These sensors observe events of interest and report them to a fusion center (FC) which may be located far from the observation area. Information processing is a major challenge to deal with in WSN applications for many reasons. Several strategies aimed at reducing communication and energy costs are available in the literature. For example, a node selection strategy based on relative geometry for localization has been proposed in [1]. Similarly, in [2], only the sensor whose measurement could provide the maximum information utility is activated at each snapshot for both system stability and tracking accuracy. There are also several low-power and sleep strategies for reducing energy costs associated with sensing and communication in WSN [3]–[6].

Our focus, in this paper, is solely on signal processing perspective to reduce the communication costs in WSN. Specifically, we look at measures that a sensor can make use of to assess the utility or value of the observations that it acquires. The benefit of such utility measures, namely test statistics here, is that in certain cases, the decision to transmit or not to transmit can be made by the sensor itself. In addition, sensors can utilize such utility measures to schedule transmission of their observations.

We begin with a deterministic source model in which a WSN consists of a number of sensors observing a phenomenon of interest and communicating their observations to a FC.

E. Ayeh, K. Namuduri, and X. Li are with the Department of Electrical Engineering, University of North Texas, Denton, TX 76207, USA (e-mails: kea0043, Kamesh.Namuduri@unt.edu).

A. Major Contributions

Our research work deals with statistical ranking of sensor test statistics for signal detection applications in WSN under the Neyman-Pearson (NP) paradigm. We consider a deterministic source observed by a network of sensors and develop a strategy for ranking the transmissions based on their test statistics. The transmitted test statistics approximate the signal-to-noise ratios (SNRs) and thus could be utilized for further analysis of the network characteristics. The proposed strategy is based on the ordered transmissions scheme presented in [7] but makes use of different statistic. We derive an upper bound on the number of sensors saved. Our result complements the lower bound derived in [7]. The main benefit of the ordered transmissions scheme is that the decision to transmit or not to transmit can be made at the sensor level which results in significant savings in transmission costs. We demonstrate these advantages through simulations and theoretical analysis.

B. Organization

Section II provides a summary of related work in the area of signal processing based data reduction strategies. Section III describes the proposed statistical ranking strategy. Section IV assesses the performance of the proposed scheme. This section includes an illustrative example to quantify the benefits of the proposed scheme. Section V concludes the paper.

II. RELATED WORK

In the following discussion, we review recent literature and discuss data reduction strategies that are closely related to our research. For example, a scheme for distributed detection based on a "censoring" or "send/no-send" idea is proposed in [8]. The sensors are assumed to "censor" their observations so that each sensor sends to the FC only "informative" observations, and leaves those deemed "uninformative" untransmitted. The problem of interest, therefore, is reduced to an N sensor binary hypothesis testing problem, where the sensors are trying to decide between the null (H_0) and alternate (H_1) hypotheses. A sensor will transmit its observation if and only if its likelihood ratio is very large or very small. The problem of energy efficient signal detection in considered in [7]. The authors proposed an ordered transmission scheme in which only sensors with the most informative observations transmit. In their approach, the i^{th} sensor will transmit after a time proportional to the inverse of its likelihood ratio $(1/|\ln(L_i)|)$ and once enough evidence is accumulated to decide for one hypothesis or the other, the process is stopped in order to save valuable transmission energy. This assume however knowledge of the a priori probabilities of the respective hypotheses, which is not possible in applications such as sonar and radar. The ordered transmission framework proposed in [7] has since been investigated by several researchers [9]–[12].

One challenge in data reduction strategies is to find the subset of significant observations before they are actually transmitted to the FC. It is possible to achieve this in certain scenarios depending on the observation model. For example, when the sensor observations are uncorrelated, each sensor can decide the value of its observations independently. In fact, in conventional cluster routing protocols, it is not necessary to involve all the sensors due to the redundancy characteristics of the information that they acquire. These observations led to the development of a data reduction strategy in which only sensors with high local SNRs are selected to transmit their observations [13].

III. Centralized Detection Based on Statistical Ranking

Consider a WSN deployed for a target detection application. It is assumed that the observed source signal is deterministic and that the observations are corrupted by an additive white Gaussian noise (AWGN). This involves at the local level, transmission of the test statistics to a FC that will decide whether or not the target is present. The goal is to develop a data reduction strategy based on statistical ranking of the test statistics, to allow only informative sensors to transmit to the FC. The detection problem at the FC thus consists of deciding between the two statistical hypotheses H_1 (target is present) and H_0 (target absent):

$$\begin{aligned} H_0 &: y_i = w_i, & 1 \leq i \leq M \\ H_1 &: y_i = A + w_i, & 1 \leq i \leq M. \end{aligned}$$

Here, $A > 0$ is the deterministic signal that is assumed known and corrupted by a zero-mean AWGN w_i with known variance σ_i^2. M represents the total number of sensors in the network. It is assumed that there is no prior knowledge about the probabilities of occurrence of H_1 and H_0.

The NP detector decides H_1 if the likelihood ratio $L(\mathbf{y})$ exceeds a threshold γ as follows

$$L(\mathbf{y}) = \frac{p(\mathbf{y}; H_1)}{p(\mathbf{y}; H_0)} > \gamma, \tag{1}$$

where $\mathbf{y} = [y_1, y_2, ..., y_M]^T$.

Starting from (1), and after some simplifications which the interested reader can find in [14], the FC decides H_1 if

$$T(\mathbf{y}) = \sum_{i=1}^{M} \frac{A y_i}{\sigma_i^2} > \gamma'. \tag{2}$$

In (2),

$$\gamma' = \sqrt{\sum_{i=1}^{M} \frac{A^2}{\sigma_i^2}} Q^{-1}\left(P_{FA}\right), \tag{3}$$

with P_{FA} the probability of false alarm, $Q(x)$ the probability that a normalized Gaussian random variable will be larger than x and $Q^{-1}(\cdot)$ is its inverse.

A. Local Measurements

Each sensor may make a set of measurements $\mathbf{x}_i = [x_i[0], x_i[1], ..., x_i[N-1]]^T$, where N is the total number of observations. After all the N measurements acquired, each sensor computes its test statistic $T_i = \frac{A y_i}{\sigma_i^2}$, where $y_i = \frac{1}{N}\sum_{n=0}^{N-1} x_i[n]$, and sends it to the FC for global decision.

B. Global Decision

The goal of the FC is to distinguish between H_1 and H_0 with $k < M$ sensor transmissions. For that, we perform a sequential test with ranked sensor test statistics. According to the ordered transmissions scheme of [7], we have $|T_1| \geq |T_2| \geq ... \geq |T_M|$. Hence, the larger the magnitude of the individual test statistic T_i, the greater the contribution of sensor i to the global test statistic $T(\mathbf{y})$. Assuming the network is synchronized, the sensors compute their test statistic at the same time but only send the result after a delay $D_i = C/|T_i|$, where C is a known normalized constant. As a result, the sensors with large test statistic values will transmit first.

When the i^{th} test statistic T_i is received at the FC, there are $M - i$ sensors that have not transmitted yet. The cumulative sum S_i of the received test statistics and R_i of the remaining $M - i$ test statistics can be expressed as shown in (4) and (5) respectively

$$S_i = \sum_{j=1}^{i} T_j = S_{i-1} + T_i, \tag{4}$$

$$R_i = \sum_{j=i+1}^{M} T_j, \tag{5}$$

with $T_j = \frac{A y_j}{\sigma_j^2}$, $i = 1, 2, ..., M$ and $S_0 = 0$.

In an unconstrained energy scenario, the FC collects all the individual test statistics, computes $T(\mathbf{y})$ and compares the result to γ' in order to make a decision. In this work, we would like to make this decision with the least number of sensor transmissions in order to save the limited resources that the network depends on. Based on the ordering scheme employed, $T(\mathbf{y})$ can be bounded using

$$S_i - (M-i)\left|T_i\right| \leq T(\mathbf{y}) = S_i + R_i \leq S_i + (M-i)\left|T_i\right|. \tag{6}$$

For the FC to declare H_1 or H_0 after collecting all the test statistics, $T(\mathbf{y})$ must be greater or less than γ' respectively. Hence, at any instant during the collection of the test statistics, if $S_i - (M-i)\left|T_i\right| \geq \gamma'$, the FC can declare H_1 and halt further transmissions. Similarly, if $S_i + (M-i)\left|T_i\right| < \gamma'$, the FC declares H_0 and halts further transmissions.

The stopping rule for determining both H_0 and H_1 can then be summarized as follows:

$$Pr\{k\} = Pr\{S_k \geq t_u(k), S_{k-1} < t_u(k-1), \cdots, S_1 < t_u(1)\} \quad (7)$$

$$Pr\{k\} = Pr\{S_k \geq t_u(k)\} * Pr\{S_{k-1} < t_u(k-1)\} * \cdots * Pr\{S_1 < t_u(1)\} \quad (8)$$

- $S_i \geq \gamma' + (M-i)|T_i| = t_u(i)$: accept H_1 and terminate the current test;
- $S_i < \gamma' - (M-i)|T_i| = t_l(i)$: accept H_0 and terminate the current test.

The overall approach to statistically ranking sensors observations can be summarized in the following steps:

1) Each sensor collects a set of measurements and estimates y_i.
2) Each sensor computes its test statistic $T_i = \frac{Ay_i}{\sigma_i^2}$.
3) Each sensor then computes $D_i = C/|T_i|$, where C is a normalized constant pre-computed and communicated to all sensors.
4) After a time period equal to D_i, sensor i will transmit T_i to the FC. As a result, the sensors with the highest test statistic will transmit first.
5) Upon receiving the test statistics, the FC computes the cumulative sum S_i and uses the stopping criteria above described to make a decision regarding the presence or absence of the target. Once one of the stopping criteria is satisfied, it broadcasts a message to halt further transmissions.

In the next section, we illustrate the benefits of the proposed strategy. We conducted different experiments and present the obtained results.

IV. Performance Evaluation

The proposed strategy will stop transmissions once one of the stopping criteria is satisfied. Under H_1, the probability that only $k < M$ sensor transmissions is required for detection is defined as shown in (7), which is then simplified to (8).

This, however, requires the knowledge of

$$Pr\{S_i \geq t_u(i)\} = \int_{t_u(i)}^{\infty} f(S_i)\, dS_i$$

and

$$Pr\{S_i < t_l(i)\} = \int_{-\infty}^{t_l(i)} f(S_i)\, dS_i, \quad (9)$$

where $f(S_i)$ denotes the probability density function (PDF) associated with the sum of i dependent test statistics $\sum_{j=1}^{i} T_j$. To determine this PDF, we need to determine $f(T_1, T_2, ..., T_i)$ first, which can be written as

$$f(T_1, ..., T_i) = f(T_1) * f(T_2|T_1) * \cdots * f(T_i|T_{i-1}), \quad (10)$$

due to the ordering of the test statistics.

There is no closed form to (10) and thus to $f(S_i)$. In fact, such closed form expressions require the distribution of a linear combination of normal order statistics with different variances, which to the best of our knowledge does not have a closed form. However, we know that a target is present in the network when

$$\sum_{i=1}^{k} T_i \geq \sqrt{\sum_{i=1}^{M} \frac{A^2}{\sigma_i^2} Q^{-1}(P_{FA})} + (M-k)|T_k| \quad (11)$$

or

$$\sum_{i=1}^{k} |T_i| + k|T_k| - M|T_k| \geq \sqrt{\sum_{i=1}^{M} \frac{A^2}{\sigma_i^2} Q^{-1}(P_{FA})}. \quad (12)$$

Let us refer to the largest and smallest test statistics after ranking as T_{max} and T_{min} respectively. Eq.(12) can be rewritten as

$$2k|T_{max}| - M|T_{min}| \geq \sqrt{\sum_{i=1}^{M} \frac{A^2}{\sigma_i^2} Q^{-1}(P_{FA})}, \quad (13)$$

which is then simplified to

$$k \geq \frac{\sqrt{\sum_{i=1}^{M} \frac{A^2}{\sigma_i^2} Q^{-1}(P_{FA})} + M|T_{min}|}{2|T_{max}|}. \quad (14)$$

Assuming N_s represents the number of sensor transmissions saved and using the fact that $k = M - N_s$, (14) is rewritten as

$$N_s \leq M - \frac{\sqrt{\sum_{i=1}^{M} \frac{A^2}{\sigma_i^2} Q^{-1}(P_{FA})} + M|T_{min}|}{2|T_{max}|}. \quad (15)$$

The bound on the number of sensor transmissions saved which depends on several factors can be obtained in a similar fashion under H_0. Since T_{max} and T_{min} are two random variables, this bound is also a random variable. Deriving the expected value for this bound requires the knowledge of the distribution of these random variables, for which closed form expressions are not known. Hence, we present our findings through several simulations.

We plot this upper bound as a function of M for different values of the signal level A and for $P_{FA} = 0.01$ in Fig. 1. The sensor observations are assumed to be corrupted by AWGN having the same mean (0) but different variances (randomly selected between 0.4 and 1).

Fig. 1 suggests that the number of sensor transmissions saved increases as the signal level increases. For example, when $A = 0.5$, the bound is 60 and when $A = 1$, the bound is slightly greater than 70. For higher values of A ($A \geq 5$), the bound asymptotically reaches 90.

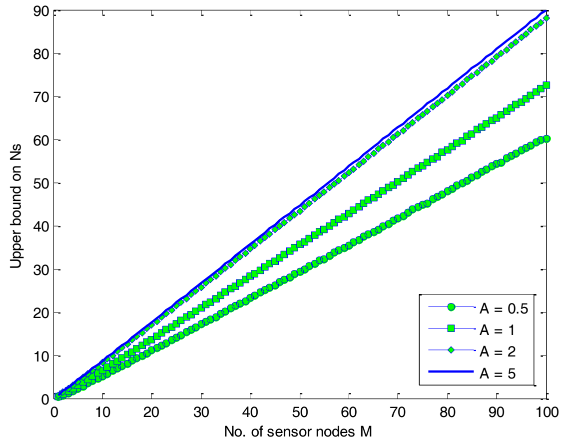

Fig. 1. Upper bound on the average number of sensor transmissions saved (N_s) over the unconstrained energy approach.

We also observe that varying P_{FA} does not have significant effect on the value of the bound.

We demonstrate the benefits of the proposed strategy through further simulations. We simulated a network of 20 nodes and a FC. The observed deterministic signal A is assumed to be the same for all the sensors, and has a value of 1. Fig. 2 shows the simulation results under H_1 for different desired P_{FA} at the FC. The results show that the proposed ranking strategy requires on average approximately 11, 11 and 10 sensor transmissions for $P_{FA} = 0.01, 0.02, 0.05$ in order to make a decision. This can be observed when $Pr\{S_i \geq t_u(i)\} = 0.5$.

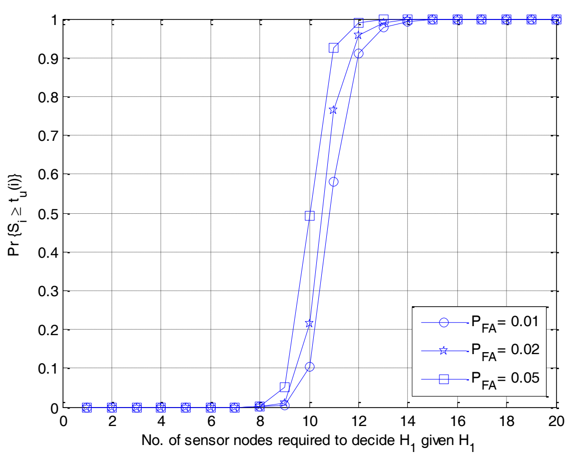

Fig. 2. Detection performance at the FC under hypothesis H_1.

Fig. 3 shows the simulation results under H_0 for different values of P_{FA} at the FC. Here, the unconstrained energy approach requires all 20 sensor transmissions in order to make a decision, whereas the proposed strategy can make this decision with less sensor transmissions since the upper bound of the global test statistic is predictable at each instant due to ranking. Thus, once the predicted global test statistic after the i^{th} transmission is less than $t_l(i)$, the FC can declare the target absent. For example, for $P_{FA} = 0.01, 0.02, 0.05$, the FC can save on average approximately 12, 11, and 10 sensor transmissions respectively, over the unconstrained energy approach. This can be observed when $Pr\{S_i < t_l(i)\} = 0.5$.

The results of Fig. 1 were obtained after 10^4 Monte Carlo realizations and those of Fig. 2 and Fig. 3 were obtained after 10^6 realizations.

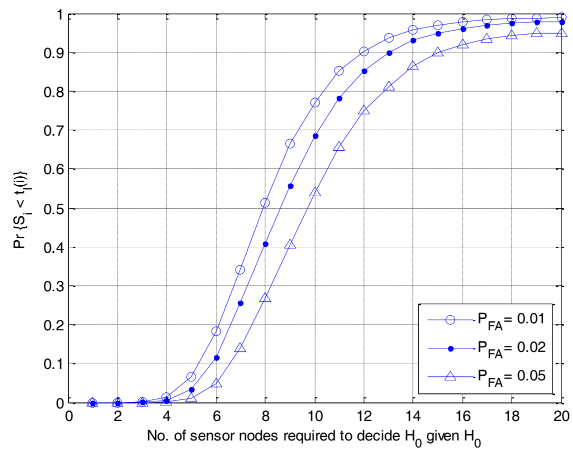

Fig. 3. Detection performance at the FC under hypothesis H_0.

V. Conclusions

This paper proposed a data reduction strategy based on statistical ranking of sensor test statistics in a centralized detection application. Based on the proposed protocol in which sensor transmissions are ordered according to the magnitude of their test statistics, each sensor can determine if and when to transmit its information to the FC. As a result, sensors with the highest test statistics transmit first while the ones with small test statistics are blocked from transmission by the FC. We demonstrated that the proposed strategy saves significant transmission costs and thus preserves limited resources that the network operates on. It has been shown through simulations that the proposed ranking strategy, under the considered scenarios, can save more than 50% of sensor transmissions when compared to the unconstrained energy approach. Theoretically, we showed that finding a tight bound on the number of required sensors is subject to determining the distribution of a linear combination of normal order statistics with different variances.

References

[1] L. Kaplan, "Global node selection for localization in a distributed sensor network," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 42, no. 1, pp. 136 – 146, January 2006.

[2] H. Ying and Z. Wei, "A novel node selection method of bearings-only sensors for target tracking in wireless sensor networks," in *Proceedings of the 2009 WRI International Conference on Communications and Mobile Computing*, vol. 01, 2009, pp. 136–140.

[3] J. Ayers, K. Mayaram, and T. Fiez, "Tradeoffs in the design of cmos receivers for low power wireless sensor networks," in *Circuits and Systems. ISCAS. IEEE International Symposium on*, May 2007, pp. 1345 –1348.

[4] J. Yang and X. Li, "Design and implementation of low-power wireless sensor networks for environmental monitoring," in *Wireless Communications, Networking and Information Security (WCNIS), IEEE International Conference on*, June 2010, pp. 593 –597.
[5] S. Sakunia, S. Bhalerao, A. Chaudhary, M. Jyotishi, M. Dixit, R. Jumade, and R. Patrikar, "Ubisens: Achieving low power wireless sensor nodes," in *Wireless and Optical Communications Networks. WOCN '07. IFIP International Conference on*, July 2007, pp. 1 –6.
[6] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization," *Mobile Computing, IEEE Transactions on*, vol. 6, no. 2, pp. 221 –236, February 2007.
[7] R. Blum and B. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *Signal Processing, IEEE Transactions on*, vol. 56, no. 7, pp. 3229 –3235, 2008.
[8] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 32, no. 2, pp. 554 –568, April 1996.
[9] R. Blum, "Ordering for estimation and optimization in energy efficient sensor networks," *Signal Processing, IEEE Transactions on*, vol. 59, no. 6, pp. 2847 –2856, June 2011.
[10] P. Braca, S. Marano, and V. Matta, "Asymptotically consistent one-bit detection in large sensor networks," *19th European Signal Processing Conference (EUSIPCO)*, pp. 1035 –1039, September 2011.
[11] ——, "Single-transmission distributed detection via order statistics," *Signal Processing, IEEE Transactions on*, vol. 60, no. 4, pp. 2042 –2048, April 2012.
[12] Y. Yang, R. Blum, and B. Sadler, "A distributed and energy-efficient framework for neyman-pearson detection of fluctuating signals in large-scale sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 7, pp. 1149 –1158, September 2010.
[13] Q. Ye, Y. Liu, L. Zhang, and C. hyun Youn, "An energy efficiency scheme using local snr for clustered wireless sensor networks," in *Mobile and Ubiquitous Systems: Networking Services, 2007. MobiQuitous 2007. Fourth Annual International Conference on*, August 2007, pp. 1 –4.
[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998.