Distributed Detection in Wireless Sensor Networks Using Dynamic Sensor Thresholds

PRIYADIP RAY AND PRAMOD K. VARSHNEY, Fellow, IEEE

Electrical Engineering and Computer Science Department Syracuse University, Syracuse, NY, USA

This paper presents a new approach for distributed target detection in wireless sensor networks (WSNs). Contrary to the conventional practice where every sensor uses an identical threshold for decision-making, an unequal and dynamic local sensor threshold selection scheme is proposed. This threshold selection scheme is based on a recently proposed statistical metric for multiple testing problems called the False Discovery Rate (FDR). Assuming a signal attenuation model, where the received signal power decays as the distance from the target increases, various performance indices like the system level probability of detection and the probability of false alarm are studied. Simulation results are provided to demonstrate the effectiveness of this approach.

Keywords Wireless Sensor Networks; Distributed Detection; Decision Fusion; False Discovery Rate

1. Introduction

Recent technological advances have enabled the deployment of a large number of networked sensors to continuously observe an event area for reliable estimation/detection/tracking of events. One particularly interesting problem for wireless sensor networks (WSNs) is distributed target detection. In distributed detection, each sensor instead of sending its raw data, sends quantized data to the fusion center to minimize power and bandwidth consumptions. The literature on distributed detection is quite extensive and numerous papers exist which deal with various aspects of distributed detection. For the conventional distributed detection problem, where multiple sensors, e.g., radars, observe a single target and receive noisy versions of the same target signal, an optimum fusion rule has been derived under the conditional independence assumption [1, 2]. Fusion rules under correlated sensor observations have been studied in [3, 4, 5].

Decision fusion for target detection in WSNs, where the received signal power decays as the distance from the target increases, has been investigated in [6, 7, 8]. In this problem, as the probability of detection of each sensor is unknown, due to unknown target location, the optimal Chair-Varshney fusion rule cannot be used. A counting rule that uses the total number of detections (1’s) reported by the local sensors has been proposed.
in these papers. In this formulation, all the local sensors use the same threshold and the system level probability of false alarm and probability of detection have been derived. The proposed scheme was also shown to be asymptotically optimal. In this paper, we break away from the equal local sensor threshold paradigm and explore the possibility of improved detection performance under unequal and dynamic sensor thresholds.

A statistical metric called the false discovery rate (FDR), and an algorithm to control FDR for multiple comparison problems was recently proposed by Benjamini and Hochberg [9]. A preliminary investigation of the application of FDR for WSNs has been conducted by Ermis and Saligrama [10,11,12]. Their work studied the application of FDR for detection and localization under communication constraints and also proposed a distributed FDR algorithm. In this paper, we propose a FDR based approach to set local sensor thresholds and study the performance of the resulting distributed target detection system.

The remainder of the paper is organized as follows. Section II discusses the target and signal models adopted. Section III introduces the problem of distributed target detection in WSNs and the concept of FDR based local sensor threshold selection scheme. Section IV presents performance analysis of FDR based sensor thresholds for distributed detection. Section V provides simulation results, and concluding remarks are provided in Section VI.

2. Problem Formulation

A total of \(N\) sensors are randomly deployed in the region of interest (ROI), which is assumed to be a square with area \(b^2\). The locations of the sensors are assumed to be unknown to the fusion center, but the position co-ordinates are i.i.d. and follow a uniform distribution:

\[
f(x_i) = \begin{cases} 
1/b & 0 \leq x_i \leq b \\ 
0 & \text{otherwise}
\end{cases}
\]  

\[
f(y_i) = \begin{cases} 
1/b & 0 \leq y_i \leq b \\ 
0 & \text{otherwise}
\end{cases}
\]

for \(i = 1, \ldots, N\) where \((x_i, y_i)\) are the coordinates of sensor \(i\).

We assume that the noises at local sensors are i.i.d. and follow the standard Gaussian distribution,

\[
n_i \sim N(0, 1)
\]

For a local sensor \(i\), the binary hypothesis testing problem is:

\[
H_1 : s_i = a_i + n_i
\]

\[
H_0 : s_i = n_i
\]

where \(s_i\) is the received signal, and \(a_i = \sqrt{P_i}\) is the signal amplitude due to the presence of target.
We assume that the signal power emitted by the target decays as the distance from the target increases. An isotropic signal power attenuation model is adopted here [8]:

\[ P_i = g(d_i) \]  \hspace{1cm} (6)

where \( P_i \) is the signal power measured at sensor \( i \) located at \((x_i, y_i)\), and \( d_i \) is the distance between the target located at \((x_t, y_t)\) and local sensor \( i \).

\[ d_i = \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2} \]  \hspace{1cm} (7)

In this paper, the following \( g(x) \) has been adopted,

\[ g(x) = \begin{cases} 
P_0 & 0 \leq x \leq d_0 \\
\frac{P_0}{x^n} & x > d_0 
\end{cases} \]  \hspace{1cm} (8)

where \( P_0 \) is the signal power measured at a reference distance \( d_0 \), \( n \) is the signal decay exponent and varies with the medium of propagation (air, water etc.) and surrounding environment (indoor, outdoor etc.). The model adopted above is very general, and can be used for modeling targets, which are the source of isotropically radiating electromagnetic or acoustic waves. The sensors are assumed to be passive sensors.

3. Distributed Target Detection

In the classic distributed detection scheme, due to bandwidth and energy constraints, each local sensor sends a hard decision, i.e., a single bit, to the fusion center. In this paper, it is assumed that the communication channels between the local sensors and the fusion center are perfect. When the fusion center has complete knowledge of the \( P_d \) and \( P_{fa} \) of each local sensor, the optimum fusion rule is given in [1, 2]. But for the problem being considered here, it is very difficult to calculate the probability of detection of the \( i^{th} \) sensor \( (P_{di}) \) as \( P_{di} \) depends on the distance between the sensor and the target. In the work by Niu and Varshney [6, 7, 8], a “counting rule” which uses the total number of detections (1’s) reported by the local sensors as a statistic, has been proposed. The “counting rule” is described below:

Let the binary data from local sensor \( i \) be denoted by \( I_i = \{0, 1\} \) where \( I_i \) takes the value 1 when there is a detection and 0 otherwise. The data from all the sensors are denoted by \( I = \{I_i: i = 1, \ldots, N\} \). After collecting data \( I \), the fusion center makes a final decision about a target’s presence or absence using the fusion rule [6, 7, 8],

\[ \bigwedge_{i=1}^{N} I_i < T \]  \hspace{1cm} (9)

where \( T \) is the system level threshold. Also, in this scheme every sensor is assumed to have the same threshold.
In this paper we depart from the equal and static local sensor threshold paradigm of [6, 7, 8], and adopt a dynamic local sensor threshold scheme. The intuition behind this approach is that as the distances between the target and the sensors are not the same, different sensors have unequal signal-to-noise (SNR) ratios; hence improved performance may be achieved by using non-identical thresholds for different sensors. We adopt a FDR based local sensor threshold selection method and assume a two-bit per sensor communication capability to implement the distributed FDR algorithm as proposed in [10,11].

3.1. False Discovery Rate (FDR)

In a breakthrough paper, Benjamini and Hochberg [9] introduced a new approach to multiple hypothesis testing that controls the false discovery rate (FDR), defined as the fraction of false rejections among those hypotheses rejected. This procedure is very appealing as it controls a metric, which is often of more practical importance than the conventional family wide error rate (FWER), an error metric popularly used when testing a family of hypotheses. Table 1 provides details regarding the definition of FWER and FDR.

Family wide error rate (FWER) is defined as the probability of committing any type 1 error, \( P(V \geq 1) \). For this problem, it is the probability of one or more sensors reporting the presence of a target (false alarm) when no target is present.

False discovery rate (FDR) is defined as the expected ratio of the number of observations falsely classified into alternate hypotheses (declared \( H_1 \) when \( H_0 \) is true) to the total number of observations classified into the alternate hypotheses (\( H_1 \)). The proportion of errors committed by falsely rejecting null hypotheses (declared \( H_1 \) when \( H_0 \) is true) can be viewed through the random variable,

\[
L = \frac{V}{V + S}
\]

(10)

FDR (\( L_e \)) is defined to be the expectation of \( L \),

\[
L_e = E(L) = E\left(\frac{V}{V + S}\right) = E\left(\frac{V}{R}\right)
\]

(11)

\( L_e = 0 \), when \( V + S = 0 \), as no false rejections can be committed.

| Table 1 |
|---------|
| Number of errors committed when testing M null hypotheses |
|          | Declared \( H_0 \) | Declared \( H_1 \) | Total |
|----------|-------------------|-------------------|-------|
| \( H_0 \) True | U | V | \( M_0 \) |
| \( H_1 \) True | T | S | \( M - M_0 \) |
| Total     | \( M - R \) | R | M |


3.2. Algorithm to Control False Discovery Rate (FDR)

Consider testing \( n \) observations based on the corresponding p-values \( P_1, P_2, \ldots, P_n \), where the p-value for an observation \( Y \) is defined as,

\[
X = p(Y) = \int_y f_0(t) \, dt
\]

where \( f_0(t) \) is the probability density function (pdf) of the observations under \( H_0 \).

The algorithm that keeps the FDR below a specified value \( q^* \), is provided below [9]:

1. Calculate the p-values of all the observations and arrange them in ascending order.
2. Find the largest index, \( i_{\text{max}} \), such that \( P(i) \leq \frac{i}{n} q^* \).
3. Declare \( P_j \) significant (i.e., all observations corresponding to these p-values are classified as \( H_1 \)) for \( 0 \leq j \leq i_{\text{max}} \).

Under independent test statistics this procedure controls the FDR at \( q^* \).

The FDR control procedure described above is a centralized procedure and requires the ordering of p-values. Ermis and Saligrama [10,11] have proposed a decentralized procedure by utilizing an extra one-bit in-network communication among the sensors. In this paper, we assume that the p-values can be ordered in a distributed manner.

4. Distributed Target Detection using FDR Based Local Sensor Threshold

In the work by Niu and Varshney [6, 7, 8] for target detection in WSNs, all the sensors employ a constant threshold \( \gamma \), which fixes the \( P_{fa} \) at each sensor to \( P_{fa} = Q(\tau) \), and the system-wide \( P_{FA} \) to:

\[
P_{FA} = \sum_{i=0}^{N} \binom{N}{i} p_{fa}^i (1 - p_{fa})^{N-i}
\]

In this problem, different local sensors will have different \( P_{fa} \), which is a function of \( d_i \). Hence under hypothesis \( H_1 \), the total number of detections no longer follows a Binomial distribution, and the system wide probability of detection \( P_D \) is difficult to derive analytically. Hence \( P_D \) may be obtained by simulations or approximated by using the Central Limit Theorem (CLT) [6, 8].

As mentioned in Section III, as different sensors have unequal SNRs, an adaptive local sensor threshold selection scheme is likely to provide performance gain. We propose control of the False Discovery Rate (FDR) for the entire family of observations (i.e. observations from the entire sensor field) to a value less than or equal to \( \gamma \). The value \( \gamma \) and the system level threshold \( T \) are related to the system-wide probability of false alarm \( P_{FA} \) which will be derived shortly. To control FDR to \( \gamma \), the local sensor thresholds have to be set following the algorithm provided earlier. Setting the thresholds in this manner is an adaptive procedure and the thresholds depend both on the number of sensors in the event region and on the observations of the local sensors, i.e., the sensor with the largest p-value is compared to \( \gamma \) and subsequent sensors are compared to \( \frac{(N-i)\gamma}{N} \).

Thus, sensors with lower p-values, i.e., higher-valued observations are compared to higher thresholds.

4.1. Performance Analysis: System Level \( P_{FA} \)

At the fusion center, the probability of a false alarm is given by,
\[ P_{FA} = Pr \left( \Delta = \sum_{i=1}^{N} I_i \geq T \mid H_0 \right) \]  

where \( \Delta = \sum_{i=1}^{N} I_i \) denotes the total number of sensors which report the presence of the target, when no target is present (false alarms) for the entire sensor field. The sensor observations under such conditions are i.i.d. and normally distributed. Hence the corresponding p-values are i.i.d. and uniformly distributed.

The probability that the graph of an empirical distribution of uniformly distributed random variables on \([0,1]\) intersects a certain straight line at a certain height for the first time is given by Dempster's formula \([13]\). For the general case, for an arbitrary number of observations, the expected number of type 1 errors has been derived by Finner and Roters \([14]\) by applying Dempster’s formula for barrier crossing distributions,

\[ P = \left( \begin{array}{c} N \\ i \end{array} \right) (1 - \gamma)^{i} \left( 1 - \frac{i}{N} \gamma \right)^{N-i-1} \]  

From Equation (14), the probability of false alarm \( P_{FA} \) at the fusion center for a system level threshold \( T \) may be shown to be,

\[ P_{FA} = \sum_{i=T}^{N} P (V = i) \]  

From Equations (15) and (16), the probability of false alarm \( P_{FA} \) at the fusion center is given by,

\[ P_{FA} = \sum_{i=T}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) (1 - \gamma)^{i} \left( 1 - \frac{i}{N} \gamma \right)^{N-i-1} \]  

which is the system-wide \( P_{FA} \) for any number of sensors.

### 4.2. Performance Analysis: System Level \( P_D \)

The exact expression for the system level probability of detection \( P_D \) is difficult to derive and will be considered in future work. In this paper, we provide simulation results for the same.

### 5. Performance Comparison

In this section, we compare the detection performance of the identical threshold approach and the FDR based approach. To compare the two detection schemes, for a specified system level threshold \( T \) and a specified system level probability of false alarm \( P_{FA} \), the threshold \( \varsigma \) (for identical threshold approach) and the FDR parameter \( \gamma \) (for FDR based approach) are obtained by numerically solving Equations (13) and (17) respectively.

In our simulation results we have chosen the following parameters: \( P_0 = 2, \ b = 4, \ d_0 = 1, \ n = 2, \ N = 2, \ N = 15 \) and a system level threshold \( T = 2 \). The simulation results are based on \( 15 \times 10^5 \) Monte Carlo runs. The sensors were randomly deployed in the event region and a target was placed at a random location in the event region during each run. The ROC for the decaying target signal model is given in Fig. 1. The ROC curves demonstrate that the proposed FDR based distributed detection framework clearly outperforms the classical distributed detection strategies.
6. Conclusion

In this paper a FDR based sensor threshold selection methodology has been proposed and its feasibility examined for distributed target detection in WSNs. The system-wide probability of false alarm for such a scheme has also been presented. The simulation results show significant improvement of detection performance for the proposed scheme over an identical threshold scheme. It thus validates the intuition that when the local sensors are observing unequal signal-to-noise ratios due to a decaying target signal model, the use of unequal and dynamic thresholds for local decision-making will provide significant benefits. In this paper we have established the potential of FDR based methods for distributed target detection. Further research on these algorithms and performance analysis is currently underway.

References

1. Z. Chair and P. K. Varshney, “Optimal data fusion in multiple sensor detection systems,” *IEEE Trans. Aerospace Electron. Sys.*, vol. 22, pp. 98–101, Jan. 1986.
2. Pramod K. Varshney, *Distributed Detection and Data Fusion*. Springer, New York, 1997.
3. P. K. Willet, P. F. Swaszek, and R. S. Blum, “The good, bad and ugly: distributed detection of a known signal in dependent Gaussian noise,” *IEEE Trans. Signal Processing*, vol. 48, pp. 3266–3279, Dec. 2000.
4. E. Drakopoulos and C. C. Lee, “Optimum multisensor fusion of correlated local decisions,” *IEEE Trans. Aerospace Electron. Sys.*, vol. 27, pp. 593–605, July 1991.
5. M. Kam, Q. Zhu and W. S. Gray, “Optimal data fusion of correlated local decisions in multiple sensor detection systems,” *IEEE Trans. Aerospace Electron. Sys.*, vol. 28, pp. 916–920, July 1992.
6. R. Niu, P. K. Varshney and Q. Cheng, “Distributed detection in a large wireless sensor network,” *International Journal on Information Fusion*, vol. 7, pp. 380–394, Dec 2006.
7. R. Niu and P. K. Varshney, “Distributed detection and fusion in a large wireless sensor network of random size,” EURASIP Journal on Wireless Communications and Networking, vol. 2005, no. 4, pp. 462–472, September 2005.

8. R. Niu and P. K. Varshney, “Performance Evaluation of Decision Fusion in Wireless Sensor Networks,” in Proceedings of the 40th Conference on Information Science and Systems, Princeton 2006.

9. Y. Benjamini and Y. Hochberg, “Controlling the false discovery rate: a practical and powerful approach to multiple testing,” J.R. Statist. Soc. B, 1995, no. 1, pp. 289–300.

10. E. B. Ermis and V. Saligrama, “Adaptive statistical sampling methods for decentralized estimation and detection of localized phenomena,” in Proceedings of ICASSP, Philadelphia, PA, March 2005.

11. E. B. Ermis and V. Saligrama, “Detection and Localization in Sensor Networks Using Distributed FDR,” in Proceedings of the 40th Conference on Information Science and Systems, Princeton 2006.

12. E. B. Ermis and V. Saligrama, “Robust distributed detection with limited range sensors,” in Proceedings of ICASSP, Honolulu, Hawaii, April 2007.

13. G. R. Shorack and J. A. Wellner, Empirical Processes with Application to Statistics. John Wiley & Sons, 1986.

14. H. Finner and M. Roters, “Multiple hypothesis testing and expected number of type 1 errors,” The Annals of Statistics, 2002, vol. 30, no.1, pp. 220–238.
