5-1989

Distributed Detection of a Signal in Generalized Gaussian Noise

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

Arif Ansari
Southern Illinois University Carbondale

[5] R. Viswanathan, S. C. A. Thomopoulos, and R. Tumuluri, "Optimal serial distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 366-376, July 1988.
[6] T. Tsitsiklis and M. Athans, "On the complexity of distributed decision problems," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 440-446, May 1985.
[7] A. R. Reibman and L. W. Nolte, "Optimal detection performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, pp. 24-30, Jan. 1987.
[8] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I. New York: Wiley, 1968.
[9] H. L. VanTrees, *Detection, Estimation and Modulation Theory*, Vol. I. New York: Wiley, 1968.
[10] J. H. Miller and J. B. Thomas, "Detectors for discrete time signals in non-Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 241-250, Mar. 1972.
[11] J. N. Tsitsiklis, "On threshold rules in decentralized detection," in *Proc. IEEE Conf. Decision Contr.*, Dec. 1986, pp. 232-236.

Distributed Detection of a Signal in Generalized Gaussian Noise

R. VISWANATHAN AND ARIF ANSARI

Abstract—**The problem of distributed detection of a signal in incompletely specified noise is considered. The noise assumed belongs to the generalized Gaussian family and the sensors in the distributed network employ the Wilcoxon test. The sensors pass the test statistics to a fusion center, where a hypothesis testing results in a decision regarding the presence or the absence of a signal. Three monotone and admissible fusion center tests are formulated. Restricted numerical evaluation over a certain parameter range of the noise distribution and the range of signal level indicates that these tests yield performances at comparable levels.**

I. Introduction

The problem of detection of a signal using a distributed network of sensors has been analyzed in the literature. In order to save transmission bandwidths, the sensors process the information they receive and pass condensed information, such as the test statistics or the decisions with regard to the presence or the absence of a signal, to the fusion center. For the best performance, it is essential that the processing at the sensors and at the fusion be optimized [1]-[9].

So far, the problem analyzed in the literature assumes a complete statistical knowledge of the received signal. However, in sonar and other underwater detection problems, the signal is embedded in a noise whose characteristics are not completely known and are changing with time. In such situations, the sensors' statistics must be based on some general characteristics of the noise density function rather than on some specific form of noise density function. In this correspondence, we consider the distributed detection of a constant signal in generalized Gaussian noise. Such a noise density function approximates physical noise encountered in different situations [10], [11].

In Section II we discuss test statistics at the sensors and at the fusion. In Section III we present the performance analysis of three different tests at the fusion center. Numerical results are shown for a three sensors network with three samples per sensor. We conclude our discussion in Section IV.

Manuscript received April 29, 1988; revised October 5, 1988. This work was supported by the SDIO/IST and managed by the Office of Naval Research under Contract N00014-86-K-0515.

The authors are with the Department of Electrical Engineering, Southern Illinois University, Carbondale, IL 62901-6603.

IEEE Log Number 8926688.

II. The Generalized Gaussian Noise and Distributed Tests

The problem of detection of a constant signal in additive noise is described by the following hypotheses testing:

$$\begin{aligned} H_0: &\quad X_j = n_j \\ H_1: &\quad X_j = n_j + \theta, \quad j \text{ an integer.} \end{aligned} \tag{1}$$

We assume that the noise n_j has a symmetric density function described by the following equation [11]:

$$f(n) = \frac{\alpha^{1/c} c}{2\Gamma(1/c)} \exp\left(-\alpha |n|^c\right). \tag{2}$$

The noise has unit variance and hence α satisfies the relation

$$\alpha^{2/c} = \Gamma(3/c)/\Gamma(1/c). \tag{3}$$

By varying the parameter c, we can control the tail of the noise density. When c equals 2 the noise reduces to the Gaussian, and for c equals 1 it becomes Laplace. In general, smaller values of c represent heavy tails. For detecting a signal in symmetric noise at a sensor, a variety of nonparametric tests such as the sign test and the Wilcoxon test exist [12]. Our choice of the Wilcoxon test is motivated by the fact that i) the Wilcoxon test is nonparametric, ii) its performance is comparable to other nonparametric tests, iii) it performs better than the sign test in most cases, and iv) the Wilcoxon statistic takes on a finite number of discrete values.

Fig. 1 shows the distributed network of sensors and the fusion center. The statistics $T_1, T_2, \cdots, T_N$ are the Wilcoxon statistics, and the test at the fusion is given as follows:

$$S(T_1, \cdots, T_N) \underset{H_0}{\overset{H_1}{\gtrless}} t. \tag{4}$$

Here S is a statistic based on $T_1, \cdots, T_N$. The observations $X_1 \cdots X_n$ at each sensor are assumed to be independent and identically distributed according to (1). Hence, the T_k's are i.i.d. A sensor performs the Wilcoxon test by ranking the absolute values of the X_i's and summing the ranks of the absolute values which are due to positive observations. The performance of the Wilcoxon test is well understood [12]. It is possible to obtain the distribution of T_k under H_0 and H_1 by enumeration. For large values of n, it is difficult to obtain the distribution. However, the mean and the variance can be found [12]:

$$E(T_k) = \sum_{i=1}^{n} i\lambda_i \tag{5}$$

$$\mathrm{Var}(T_k) = \sum_{i=1}^{n} i^2 \lambda_i (1 - \lambda_i) \tag{6}$$

$$\lambda_i = N \binom{N-1}{i-1} \int_0^\infty \left[F(u) - F(-u)\right]^{i-1} \cdot \left[1 - F(u) + F(-u)\right]^{n-i} f(u)\, du \tag{7}$$

where $f(\ )$ is the density of the observation X_i and $F(\ )$ is the corresponding CDF.

We consider three different statistics at the fusion. The minimum test is given by the rule

$$\mathrm{Min}\,[T_1, \cdots, T_N] \underset{H_0}{\overset{H_1}{\gtrless}} t_m \tag{8}$$

where t_m is chosen to obtain a specific false alarm probability at the fusion center. However, when T_k's given the hypothesis are i.i.d., if any order statistic of $\{T_k\text{'s}\}$ is used as a test statistic at the fu-

0096-3518/89/0500-0775$01.00 © 1989 IEEE

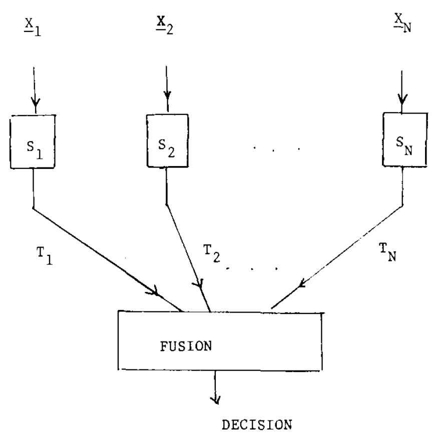

Fig. 1. Distributed sensor network.

sion, then the same performance can also be achieved by allowing the sensors to make decisions based on T_k's and by using an appropriate counting rule based on their decisions. To see this, consider the rules based on the lth-order statistic, $T_{[l]} \gtrless_{H_0}^{H_1} t$, and the counting rule based on the sensors tests $T_k \gtrless_{H_0}^{H_1} t$, and the fusion rule which declares the signal present if at least $(N - l + 1)$ sensors decide the presence of a signal. The probabilities of detection at a fixed false alarm probability for these two fusion tests are identical. The probability of detection is given by the following expression:

$$P_D = 1 - \sum_{j=l}^{N} \binom{N}{j} F_1^j(t)\left(1 - F_1(t)\right)^{N-j} \tag{9}$$

where $F_1(\ )$ is the CDF of X_j in (1) under H_1. Therefore, the full benefit of transmitting the statistic instead of the decision may be lost in this combining procedure at the fusion. Numerical performance analysis of the minimum test is presented in the next section. The tests based on other order statistics are not considered because, without randomization, they achieve only large false alarm probabilities. The next test, termed the linear Wilcoxon, is based on the sum of the T_k's.

$$\sum_{k=1}^{N} T_k \gtrless_{H_0}^{H_1} t_l. \tag{10}$$

Finally, for small n and N, a symmetric fusion test given by the following rule is considered:

$$\text{Reject } H_0 \text{ for } \{T_k \geq t_j, \text{ all possible combinations of } k \text{ and } j\}. \tag{11}$$

We consider several of these ad hoc tests at the fusion because, in general, no uniformly most powerful test based on the $\{T_k\text{'s}\}$ exists. All the three tests considered are "reasonable" in the sense that they are monotone, i.e., if certain $\{T_k^*\}$ is in the reject H_0 region, then $\{T_k \geq T_k^*\}$ is also in the rejection region. In Section IV these are all shown to be admissible. One can design best tests at the sensors and at the fusion only when the statistical distribution of the observation is completely known.

III. Numerical Performance Analysis

In order to show the small sample performance, we consider a network of three sensors $(N = 3)$ and a sample size of $3 (n = 3)$. The Wilcoxon test statistic T_k assumes values 0–6. Under H_0, T_k is distributed with $\Pr(T_k = m) = 1/8$, $m \neq 3$, and $\Pr(T_k = 3) = 2/8$. Under H_1, the probability of T_k is given by the following sum:

$$\sum \prod_{i=1}^{l} \lambda_{r_i} \prod_{1 \leq j \neq r_i \leq N} (1 - \lambda_j) \tag{12}$$

where λ_i's are given in (7), exactly l specified ranks $r_1, r_2, \cdots, r_l$ out of n have positive signs, and the summation is extended over all sets of assignments of positive signs which will lead to that value of T_k. Since we consider only small values of N, the distributions of the linear Wilcoxon, the minimum, and the symmetric test can easily be obtained once the distributions of T_k under H_0 and H_1 are calculated. For example, in the case of linear Wilcoxon, the discrete convolution is employed twice. To avoid any heavy randomization, attention is restricted to only nonrandomized tests at the fusion center. Figs. 2–5 show the performance of the three tests. For all the tests, for weak signal (small values of θ), the detection power is larger for heavy tail noise and the converse is true for strong signal. By looking at the trends in these figures, we observe that all the tests perform comparably well. Since the probabilities of false alarms do not match with nonrandomized testing, actual comparison of the tests is not possible. In Fig. 6 we show the performances of linear detector (linear detector computes the sum of all the observations at all the sensors and compares the sum to a threshold), which is optimum for Gaussian, and the linear Wilcoxon detector. For moderate signal strength, the loss associated with the linear Wilcoxon as compared to linear detector is clearly seen.

For large N, the performances of any two tests could be compared by Asymptotic Relative Efficiency (ARE). It is given by the ratio of the sample sizes required by the two tests to achieve the same detection probability and false alarm probability as the signal level goes to zero and as both the sample sizes tend to infinity [12]. By using the approach in [12], the ARE of the linear Wilcoxon test with respect to the linear detector can be computed to yield

$$\text{ARE} = \frac{6\left(1 + \dfrac{n-1}{2^{1/c}}\right)^2 \alpha^{2/c}}{(n+1)(2n+1)\Gamma^2(1 + 1/c)}. \tag{13}$$

The above ARE is plotted in Fig. 7 for various c values. For heavy tailed noise, the linear Wilcoxon has ARE larger than 1 compared to the linear detector.

IV. Discussion

The problem of detection of a constant signal in incompletely known noise with a distributed network of sensors is considered. The noise assumed is symmetric and has a generalized Gaussian density function. The sensors employ the Wilcoxon test and pass the test statistics to a fusion center. Since the statistics of the noise are not completely known, there exists no uniformly best test at the fusion center. We consider three monotone tests based on the Wilcoxon test statistics.

Among the tests considered for fusion, choosing a test which performs better than the rest is dependent on the signal level, the parameter c, and the operating false alarm probability. All possible tests, which are reasonable in the sense that they are monotone, would then have to be enumerated. However, the monotonicity condition is satisfied by so many possible methods of combination, a search over all possible tests of this class is quite complex. A combined test procedure is said to be admissible if it provides a (not necessarily the only) most powerful test against some alternative hypothesis. It may seem reasonable to narrow our search to tests that are admissible. Unfortunately, the class of admissible combined test procedures is still quite large [13]. From Table I it is seen that all the tests discussed are admissible. However, without

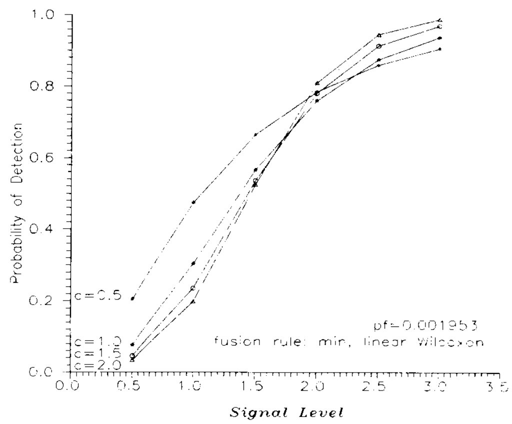

Fig. 2. Signal level versus probability of detection for the linear Wilcoxon rule with $pf = 0.001953$.

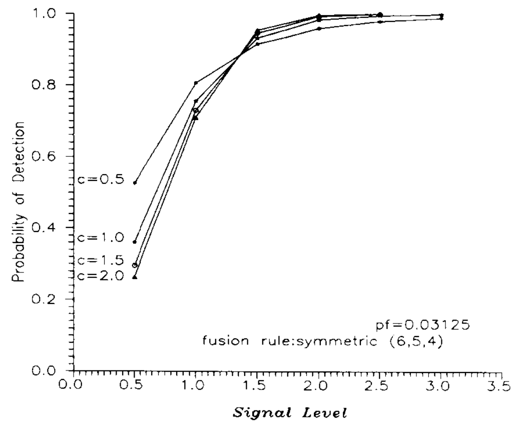

Fig. 5. Signal level versus probability of detection for symmetric (6, 5, 4) fusion rule with $pf = 0.03125$.

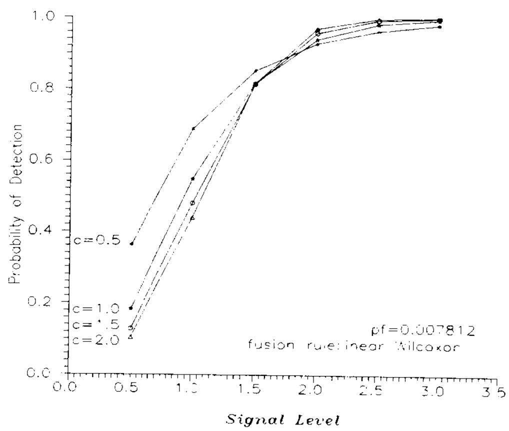

Fig. 3. Signal level versus probability of detection for the linear Wilcoxon rule with $pf = 0.007812$.

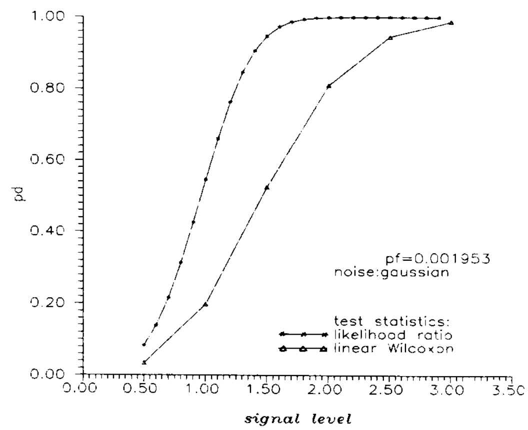

Fig. 6. Signal level versus probability of detection for the linear Wilcoxon test and likelihood ratio test with $pf = 0.001953$.

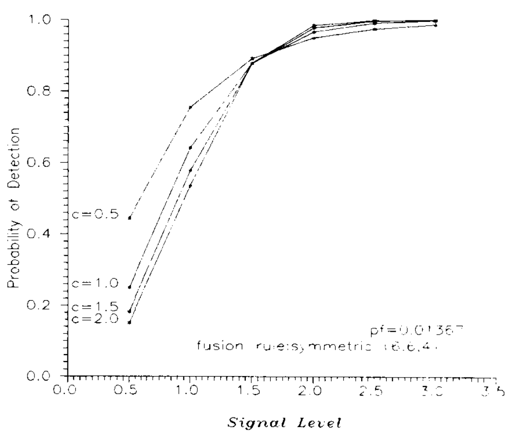

Fig. 4. Signal level versus probability of detection for symmetric (6, 6, 4) fusion rule with $pf = 0.01367$.

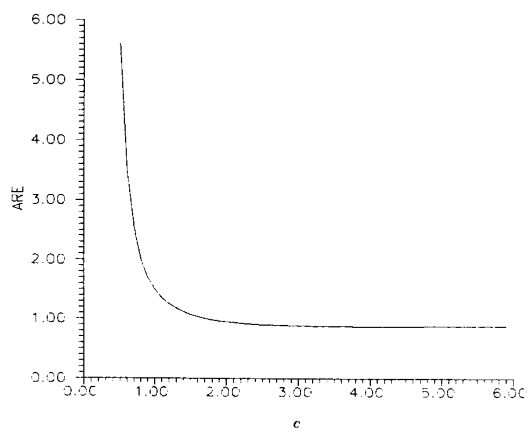

Fig. 7. ARE of linear Wilcoxon compared to linear versus c.

TABLE I
OPTIMUM LIKELIHOOD RATIO TEST AND EQUIVALENT TESTS

Signal Level	Likelihood Ratio Threshold	Equivalent Tests Threshold in parenthesis
1.0	100	$c\varepsilon[0.5,2.0]$ Linear Wilcoxon (18) Same as Min (18) or Symmetric (6,6,6)
1.0	36	$c\varepsilon[0.5,2.0]$ Linear Wilcoxon (17) Same as Symmetric (6,6,5)
1.0	11.39	$c\varepsilon[0.5,1.0]$ Symmetric (6,6,4) $c\varepsilon[0.5,2.0]$ Linear Wilcoxon (16)
2.0	2.55	$c\varepsilon[1.0,2.0]$ Symmetric (6,6,4)

some knowledge of signal level and the noise parameter c, an optimum choice of a particular test from the class of monotone admissible tests does not seem possible.

REFERENCES

[1] R. R. Tenney and N. R. Sandell, Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501-510, July 1981.

[2] R. Srinivasan, "Distributed radar detection theory," *IEE Proc.*, vol. 133, Pt. F, no. 1, pp. 55-60, Feb. 1986.

[3] J. J. Chao and C. C. Lee, "A distributed detection scheme based on soft local decisions," in *Proc. 24th Annu. Allerton Conf. Commun.*, Contr., pp. 974-993.

[4] L. K. Ekchian and R. R. Tenney, "Detection networks," in *Proc. 21st IEEE Conf. Decision Contr.*, Dec. 1982, pp. 686-690.

[5] S. C. A. Thomopoulos, R. Viswanathan, and D. P. Bougoulias, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 644-653, Sept. 1987.

[6] R. Viswanathan, S. C. A. Thomopoulos, and R. Tumuluri, "Optimal serial distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 366-376, July 1988.

[7] F. A. Sadjadi, "Hypotheses testing in a distributed environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 134-137, Mar. 1986.

[8] A. R. Reibman and L. W. Nolte, "Optimal detection performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 24-30, Jan. 1987.

[9] J. Tsitsiklis and M. Athans, "On the complexity of distributed decision problems," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 440-446, May 1985.

[10] V. R. Algazi and R. M. Lerner, "Binary detection in white non-Gaussian noise," M.I.T. Lincoln Lab, Lexington, MA, Tech. Rep. DS-2138.

[11] J. H. Miller and J. B. Thomas, "Detectors for discrete time signals in non-Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 241-250, Mar. 1972.

[12] J. D. Gibbons, *Nonparametric Statistical Inference*. New York: McGraw-Hill, 1971.

[13] V. Hedges and I. Olkin, *Statistical Methods for Meta-Analysis*. New York: Academic, 1985.

A Block Coding Technique for Encoding Sparse Binary Patterns

GENGSHENG ZENG AND NASIR AHMED

Abstract—**This correspondence introduces an efficient method for encoding sparse binary patterns. This method is very simple to implement and performs in a near optimum way.**

Manuscript received April 2, 1988; revised October 14, 1988. This work was supported by the Sandia National Laboratories, Albuquerque, NM 87185.

The authors are with the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131.

IEEE Log Number 8926691.

I. INTRODUCTION

The purpose of this correspondence is to introduce an efficient method for encoding *sparse binary patterns (images)*, where the term "sparse" implies that the patterns consist of a small number of ones, relative to the number of zeros.

The technique we consider will be referred to as *block coding*. It is shown that block coding enables us to encode sparse binary patterns with *average code word lengths* $L_{av}(p)$ that compare very closely to the *source entropy* $H(p)$ when p is small, where p is the probability of finding a one in the given pattern. Since $L_{av}(p)$ closely approximates $H(p)$, we can view such block codes as being close to optimum for encoding sparse binary patterns [1].

The sparse pattern we deal with is assumed to be a memoryless binary source. This kind of pattern is found in a 3-D authentication scheme [2]. In data compression, the patterns are usually not memoryless sources. However, when LPC (Linear Prediction Coding) is applied, the resulting error pattern is very close to a memoryless model. Yasuda [3] presented some effective methods to decorrelate 2-D facsimile patterns. For example, Boolean algebra prediction functions [3, p. 834] are shown to be very effective for typical typewritten English and Japanese documents and weather maps. The error patterns that result via prediction are sparse, and hence our block coding technique may be useful for this application also.

After the block coding method is introduced, it is compared to some other existing methods.

II. BLOCK CODING

For the purposes of discussion, we consider a (128×128) sparse binary pattern in which the probability of finding a one is $p = 0.01$. As such, the probability of finding a zero is $q = 1 - p = 0.99$. If this pattern is scanned on a row-by-row basis, it follows that we obtain a 1-dimensional array consisting of $n = 16\,384$ bits.

The proposed block coding scheme consists of the following steps.

1) Map a 2-D image into a 1-dimensional array by row-by-row scanning. The 1-dimensional array consists of $n = 2^M$ ones and zeros.

2) Divide the 2^M bit-string obtained in Step 1) into 2^a blocks, with each block consisting of 2^b bits; it then follows that $M = a + b$.

3) Between any two adjacent blocks we introduce a *comma*, which is encoded as a "0."

4a) If there is no one in a block, then no coding is needed for the block.

4b) If there are ones in a block, then assign each one a prefix "1" followed by b bits to indicate its location in the block. This location is with respect to the left end of the 2^b-bit string and numbered from 0 through $2^b - 1$. The reason for the prefix "1" is to realize an instantaneous code.

5) The bit-string resulting from Step 4) is desired code.

The above steps are best illustrated by a simple example. Suppose we consider the (4×4) binary pattern in Fig. 1. Then the results obtained via Steps 1)–5) above are as summarized in Fig. 1, for the case $M = 4$, and $a = b = 2$.

The decoding procedure is just the reverse of the coding procedure.

In general, if we have k ones and $n - k$ zeros, the code length, $L(k)$, is given by

$$\begin{aligned} L(k) &= (\text{rows} - 1) + k(\text{coding bits per one}) \\ &= (2^a - 1) + k(b + 1) = (2^a - 1) + k(M - a + 1). \end{aligned} \tag{1}$$

The probability of k ones and $n - k$ zeros occurring is

$$\Pr(k) = \binom{n}{k} p^k q^{n-k}. \tag{2}$$