An Algorithm for Processing Data Received from a Distributed Monitoring Network Consisting of Several Sensors

E Chernetsova¹, A Shishkin²

¹Faculty of Information Systems and Geotechnology, Russian State Hydrometeorology University, Voronezh st. 79, St-Peterburg, 192007, Russia
²Faculty of Information Systems and Geotechnology, Russian State Hydrometeorology University, Voronezh st. 79, St-Peterburg, 192007, Russia

E-mail: chernetsova@list.ru

Abstract. The paper presents an algorithm for detecting an object during the coherent reception of signals coming from an environmental monitoring network consisting of several sensors. The sensors of the monitoring network can change their location in space, for example, as it happens in a freely drifting system of sensors for detecting oil pollution on the water surface. The proposed algorithm uses statistics that take into account the most stable features of the distribution of the source data. It provides a constant probability of false alarm at any noise level. This algorithm can be implemented in software in an automated decision support system for the presence or absence of environmental pollution. At the same time, decisions on the detection of a monitoring object made by an automated system will be more reliable.

1. Introduction

To carry out environmental monitoring it is necessary to conduct continuous observations over time, based on a well-considered distribution of measuring instruments in space, for which it is necessary to use a stationary distributed multi-sensor remote monitoring system [1]. It should work quickly, preferably in real time. A stationary network of stations included in the monitoring system requires the availability of communication channels with a monitoring and control point (MCP) [2]. Laying a cable communication network is often unprofitable. Therefore, for communication purposes it is necessary to use a radio channel or satellite communication [3]. Since the sensors of the monitoring network receive energy from the batteries, in order to save energy in the monitoring network, it is often justified not to pre-process the signal on the sensor, but to send analog signals to the MCP, which is charged with processing the sensor signals and detecting the monitoring object [4]. Information exchange over the radio channel raises the problem of detecting an analog signal with an unknown law of fluctuations against the background of noise with an unknown distribution [5]. To solve this problem, in this paper, it is proposed to develop an algorithm for detecting a monitoring object during the coherent reception of signals from a monitoring network consisting of several sensors.
2. Theoretical analysis

Let us consider the problem of coherent detection of a signal from an object distributed in N resolution elements, which are sensors of a monitoring network. It was shown in [6] that the optimal detector is that which calculates the likelihood ratio:

\[
I(X) = \frac{1}{N} \left( \frac{\gamma_0}{\gamma_1} \right)^k \prod_{n=1}^{N} \exp \left[ \frac{1}{2\gamma_0} \sum_{n=1}^{k} x_n - \frac{1}{2\gamma_1} \sum_{n=k+1}^{N} x_n \right]^2
\]

where \( x_n \) - detector output envelope samples, \( n = 1,2, \ldots, N \)

\( \gamma_0 \) and \( \gamma_1 \) - signal variances received from \( (N-k) \) sensors, that did not fix the object and \( k \) sensors, fixed object accordingly.

From equation (1) we can see, that that detector which is optimal accordingly to the signal-to-noise ratio criterion can be implemented by a rather complex circuit, and, in addition, for its implementation a priori information is required about the parameters of signal \( (\gamma_1) \) and noise \( (\gamma_0) \), which, as a rule, in real monitoring conditions are unknown. Therefore the rule (1) characterizes the potential for detecting an object and cannot be realized in many practical cases.

It is necessary to develop an optimal by signal / noise criterion algorithm for coherent detection of a signal from a monitoring object received from \( (N-k) \) sensors on the background of noise interference provided that the signal and noise parameters, as well as the position of the fixed object \( k \) sensors among \( N \) sensors of the monitoring network are a priori unknown. Detection is formulated as the statistical task of testing general linear hypotheses [7-14] and the optimal rule is found in the class of so-called invariant rules [15].

We use the following premises [16]:

1. There are statistically independent radio pulses sent by \( N(N >> 1) \) sensors. In the absence of the object of observation, these pulses have the same average power. The law of the distribution of the noise background is considered normal.
2. In the presence of an object of observation, the resulting fluctuation in resolution is the additive sum of the signal with unknown amplitude \( \xi_m (m = 1,2, \ldots, k) \) and Gaussian noise with unknown variance \( \sigma^2 \). Coherent processing is assumed. Independent voltage samples are taken at the output of the linear path of the MCP receiver at time instants following the resolution interval. \( x_n (n = 1,2, \ldots, N) \).
3. Processing is carried out during the \( p \) periods of the signal, so that each reference element \( n \) will correspond to a sample vector \( (x_{n1}, \ldots, x_{np}) \) with multidimensional normal probability density

\[
g(X_n) = \frac{\sqrt{|A|}}{(2\pi)^{p/2}} \exp \left[ -\frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij} (x_{ni} - \xi_{ni})(x_{nj} - \xi_{nj}) \right].
\]

The mean values and the covariance matrix of the vector are determined from the expressions

\[
E(x_{ni}) = \xi_{ni}; \quad E(x_{ni} - \xi_{ni})(x_{nj} - \xi_{nj}) = \sigma_{ij}; \quad \sigma_{ij} = A^{-1}, \quad \text{where } E - \text{ is the sign of mathematical averaging, and } \xi_n = 0, \text{ if } n \in (N-k), \text{ and } \xi_n > 0 \text{ at } n \in k. \text{ It is also believed that the matrix } A = (a_{ij}) - \text{ is common to all vectors } N, \text{ having dimension } p, \text{ but unknown.}
\]

The challenge is that by sample
Matrix \( X \) consists of \( p \) column vectors \((x_{i1}, \ldots, x_{in_i})\), and each such vector has its own mean value vector \( \overline{\xi}_i = (\xi_{i1}, \ldots, \xi_{in_i}) \).

Given the accepted assumptions, the task of detection is to test complex hypotheses \( H_0 \) and \( H_1 \) regarding parameters \( \overline{\xi}_i \) and \( \Delta \).

Hypothesis testing (3) fits into the scheme of testing multidimensional linear hypotheses. As follows from the general theory [17], principles of invariance and sufficiency allows you to reduce the sample \( X \) when testing hypotheses (3) to maximally invariant statistics of the form

\[
T = \sum_{i=1}^{p} \sum_{j=1}^{p} \frac{N x_i x_j}{\sum_{n=1}^{N} (x_{ni} - \overline{x}_i)(x_{nj} - \overline{x}_j)}
\]

and the set of parameters \( \overline{\xi}_i \) and \( (a_{ij}) \) - to maximal invariant

\[
\delta^2 = N \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij} \eta_i \eta_j
\]

In expressions (4), (5) \( X_{i(j)} = N^{-1} \sum_{n=1}^{N} x_{ni(j)} \); \( \eta_{i(j)} = E(x_{i(j)}) \). Numerator of the formula (4) has off center \( \chi_p^2 \) - off-center distribution with the noncentrality parameter \( \delta^2 \) and \( p \) degrees of freedom, and the denominator has central distribution \( \chi^{2(N-p)} \), so the statistics \( (N-p)T/p \) has off-central \( F \) distribution with \( p \) and \( (N-p) \) degrees of freedom and with the noncentrality parameter \( \delta^2 \).

Regarding the parameter \( \delta^2 \) of \( F \) - distribution initial hypotheses (3) can now be formulated as follows:

\[
H_0 : \delta = 0; \quad H_1 : \delta > 0
\]

Using the method of constructing optimal rules [18], it can be shown that the most powerful invariant criterion for testing hypotheses (6) has a critical region of the form

\[
T > C.
\]

Threshold level \( C \) determined by the given probability of false alarm \( \alpha \) from the condition

\[
3
\]
\[
\int_{C}^{\infty} F_{p,(N-p)}(y)dy = \alpha
\]  

(8)

where \( F_{p,(N-p)} \) is central \( F \) distribution with \( p \) and \( (N-p) \) degrees of freedom.

The expressions (4), (7) determine the functional scheme of the detector with completely unknown correlation properties of vectors \( (x_{n1}, \ldots, x_{np}) \). For practical implementation, algorithm (7), (8) can be concretized, for example, in the case of the absence of inter-period correlation. In this case, the discovery rule and parameter \( \delta^2 \) take the form

\[
\sum_{i=1}^{p} \frac{N x_i^2}{\sum (x_{ni} - \bar{x}_i)^2} > C,
\]

(9)

\[
\delta^2 = \sum_{i=1}^{p} \bar{q}_i,
\]

(10)

where \( \bar{q}_i = \left( \sum_{n=1}^{N} \varepsilon_n \right)_i / N \sigma^2 \) - is the average for all \( N \) signal-to-noise ratio for one observation period. The detector efficiency is determined by the power function of rule (7), (8), which shows the dependence of the probability of correct detection on the parameter \( \delta^2 \). It can be calculated directly from off-center tables of \( F \) - distribution [19].

### 3. Results

Figure 1 shows the curves characterizing the effectiveness of the detector of oil pollution of the water surface depending on the resolving power of the network of contact sensors constructed in accordance with the algorithm described by expressions (4), (7). Characteristics calculated for false alarm probability value and the number of received signal periods provided that the value of the signal-to-noise ratio averaged over all \( N \) sensors for one observation period is independent of resolution (uniform distribution of translational buoys (contact sensors) along the length of contamination). For comparison, the same figure shows the power function of the potential most powerful rule (MP) of coherent detection of a known signal [20-21] in the presence of only one sensor ( ).

![Figure 1](image_url)

**Figure 1.** The probability of detecting oil pollution of the water surface by signals received from a network of contact sensors.
It can be seen from the figure 1 that ignorance of the noise and signal levels in the decision elements leads to losses in the signal-to-noise ratio. However, with increasing resolution, the detector’s efficiency increases. This is due to the fact that the increase allows a more accurate assessment of noise and signal levels. So, when the loss in the signal-to-noise ratio is ~4 dB, and when less than 1 dB.

4. Conclusion
The proposed algorithm for processing data received from a distributed monitoring network consisting of several sensors has the following practically important properties: a) does not depend on a priori unknown parameters $\sigma^2$ and $\xi_n$ ($n = 1, 2, \ldots, N$) and provides a constant probability of false alarm at any noise level; b) is invariant to the location of k sensors that fixed the object and ($N - k$) sensors that have not fixed the object, among N sensors of the monitoring network; c) has the highest probability of correct detection, depending on the average signal-to-noise ratio and for large $N > p$ close to potential.

The practical significance of the results is the development of an algorithm that is resistant to changes in the signal-to-noise ratio in the communication channels of the sensors of the monitoring network with a monitoring and control post. The algorithm can be implemented programmatically using various programming languages and used to automate the process of classification of monitoring objects at a monitoring and control point.

5. References
[1] Krapivin V & Shutko A 2012 Information Technologies for Remote Monitoring of the Environment 1st edn Springer, Berlin Heidelberg
[2] Chernetsova E A, Shishkin A D 2019 The principles of information integration in an integrated system of remote monitoring pp 126-130 in the Digest of articles “Information Technologies and Systems: Management, Economics, Transport, Law” Issue 1(33) Ed. Doctor of Technical Sciences, prof. Istomin E.P. (St. Petersburg: Andreevsky Publishing House LLC) 199 p
[3] Brooks R R, Ramanathan P, Sayeed A 2002 Distributed target tracking and classification in sensor networks IEEE Signal Processing Magazine vol 19 2 pp 17-29
[4] Wu and J Hu 2010 Design and Implementation of Production Environment Monitoring System Based on GPRS-Internet In 4th International Conference on Genetic and Evolutionary Computing pp 818-821 (Shenzhen)
[5] Gabriel Nallathambi, Jose C 2019 Principe Theory and Algorithms for Pulse Signal Processing CoRR abs/1901.01140
[6] Dan L, Wong K, Hu Y, Sayeed A 2002 Detection, classification and tracking of targets in distributed sensor networks IEEE Signal Processing Magazine vol 19 2 pp 17-29
[7] Venkatasubramaniam P, Adireddy S, Tong L 2004 Sensor networks with mobile access: Optimal random access and coding IEEE J.Sel.Areas Commun (Special Issue on Sensor Networks) vol 22 pp 1058-1068
[8] Lei Song, Hongchang Hu and Xiaosheng Cheng 2012 Hypothesis Testing in Generalized Linear Models with Functional Coefficient Autoregressive Processes Hindawi Publishing Corporation Mathematical Problems in Engineering pp 2-18
[9] Azrak R and Mélard G 2006 Asymptotic properties of quasi-maximum likelihood estimators for ARMA models with time-dependent coefficients Statistical Inference for Stochastic Processes vol 9 3 pp 279–330
[10] Maller R A 2003 Asymptotics of regressions with stationary and nonstationary residuals Stochastic Processes and Their Applications vol 105 1 pp 33–67
[11] Bai Y, Fung W K and Zhu Z 2010 Weighted empirical likelihood for generalized linear models with longitudinal data Journal of Statistical Planning and Inference vol 140 11 pp 3446–3456
[12] Fahrmeir L and Kaufmann H 1985 Consistency and asymptotic normality of the maximum
likelihood estimator in generalized linear models *The Annals of Statistics* vol 13 1 pp 342–368

[13] Carsoule F and Franses P H 2003 A note on monitoring time-varying parameters in an autoregression *International Journal for Theoretical and Applied Statistics* vol. 57 1 pp 51–62

[14] Logothetis A, Isaksson A 1999 On sensor scheduling via information theoretic criteria (in Proc. Amer. Control Conf., San Diego) CA pp 2402-2406

[15] Fuller W A 1996 Introduction to Statistical Time Series, John Wiley & Sons (New York, NY, USA, 2nd edition)

[16] Hamilton J D 1994 Time Series Analysis, 1st edn (Princeton University Press, Princeton, NJ, USA)

[17] Sergio Albeverio, Raphael Hoegh-Krohn, Jens Erik Fenstad, and Tom Lindstrom 1990 Nonstandard methods in stochastic analysis and mathematical physics, 1st edn , Academic Press, Orlando

[18] Zhao F, Liu J, Liu J, Guibas L, and Reich J 2003 Collaborative signal and information processing: An information directed approach *Proceedings of the IEEE* vol. 91 8 pp 199–1209

[19] The F Distribution and the F-Ratio https://openstax.org/books/introductory-statistics/pages/13-2-the-f-distribution-and-the-f-ratio, last accessed 2020/01/21

[20] 1990 Data communication systems and their performance :proceedings of the IFIP TC6 Fourth International Conference on Data Communication Systems and Their Performance, Barcel-iona, Spain

[21] Gimaltdinov I K, Levina T M, Stolpovskii M V, Solovev D B 2018 Dynamics of the Localized Pulse in Bubbly Liquid *IOP Conference Series: Materials Science and Engineering* 463 Part 1, Paper № 022002. [Online]. Available: https://doi.org/10.1088/1757-899X/463/2/022002