KYBERNETIKA — VOLUME 18 (1982), NUMBER 3

ON THE DETECTION OF A FINITE BINARY SEQUENCE IN THE PRESENCE OF AN INTERFERING BINARY SEQUENCE

LUDVÍK PROUZA

The problem of detecting a binary sequence in the presence of another binary interfering sequence is investigated using the least squares method. The accent is posed on finding useful detection thresholds.

1. INTRODUCTION

Recently, some interest is concentrated on the problem of the detection of discrete signals in the presence of noise and of other interfering discrete signals.

Both more heuristic and more theoretic approaches can be found in the literature, [1], [2] being typical. There seems that practically usable solutions are based on the supposition that the interfering signal is substantially stronger than the wanted one and than the noise.

In what follows, the usefulness of a simple approach based on this supposition and using the method of least squares will be investigated for signals of the form of finite binary sequences.

2. SOME FUNDAMENTAL RELATIONS

Let $N > 1$ denote the number of terms of all used sequences. Let

(1) $$\{s_1, s_2, \ldots, s_N\}, \quad s_i = S \operatorname{sign} s_i \quad (i = 1, \ldots, N)$$

be the wanted sequence, let

(2) $$\{u_1, u_2, \ldots, u_N\}, \quad u_i = U \operatorname{sign} u_i, \quad (i = 1, \ldots, N)$$

be the interfering sequence, let

(3) $$\{n_1, n_2, \ldots, n_N\},$$

be the noise sequence with independent Gaussian terms, all $\mathsf{N}(0, \sigma)$.

The received sequence will be

$$\{x_1, x_2, \ldots, x_N\}\,, \quad x_i = s_i + u_i + n_i\,, \quad (i = 1, \ldots, N)\,. \tag{4}$$

Thus, for the sake of simplicity, the time coincidence of (1), (2), (3) is supposed. S, U of (1), (2) are not known, moreover, there will be supposed

$$U \gg S\,, \quad U \gg \sigma\,, \tag{5}$$

so that with practical certainty

$$\operatorname{sign} x_i = \operatorname{sign} u_i\,. \tag{6}$$

In the place of S, U, one will seek s, u fulfilling

$$\sum_{i=1}^{N} (x_i - s\beta_i - u \operatorname{sign} x_i)^2 = f(s, u) = \min\,. \tag{7}$$

where

$$\{\beta_1, \beta_2, \ldots, \beta_N\}\,, \quad \beta_i = \pm 1\,, \quad (i = 1, \ldots, N) \tag{8}$$

is a given binary sequence.

From (6), there follows

$$\partial \operatorname{sign} x_i / \partial s = 0\,, \quad (i = 1, \ldots, N) \tag{9}$$

and from (7), the system of necessary conditions will be obtained:

$$\begin{aligned} Cs + Nu - \sum |x_i| &= 0\,, \\ Ns + Cu - \sum x_i \beta_i &= 0\,. \end{aligned} \tag{10}$$

In (10) and in what follows, both summing bounds are the same as in (7) and will be omitted. Further

$$C = \sum \beta_i \operatorname{sign} x_i = \sum \beta_i \operatorname{sign} u_i \tag{11}$$

(the right side with respect to (6)).

The solution of (10) is

$$s = \frac{N \sum \beta_i x_i - C \sum |x_i|}{N^2 - C^2}\,, \tag{12}$$

$$u = \frac{N \sum |x_i| - C \sum \beta_i x_i}{N^2 - C^2}\,. \tag{13}$$

By geometric consideration, it is clear that this solution makes (7) to the minimum. Clearly, s, u are random variables and for the detection of the wanted signal, s is of primary importance if (5) holds.

Firstly, the numerator of (12) will be investigated. Denoting

$$Nx_i - \left(\sum |x_i|\right) \operatorname{sign} x_i = \zeta_i\,, \quad (i = 1, \ldots, N)\,, \tag{14}$$

it is easily obtained

(15) $$\sum \beta_i \zeta_i = \sum (N\beta_i - C \operatorname{sign} u_i)(s_i + n_i)$$

and (denoting E the expected value)

(16) $$\mathsf{E}(\sum \beta_i \zeta_i) = \sum (N\beta_i - C \operatorname{sign} u_i) s_i = S(ND - CH),$$

where

(17) $$D = \sum \beta_i \operatorname{sign} s_i ,$$

(18) $$H = \sum \operatorname{sign} s_i \operatorname{sign} u_i .$$

From (15), (16)

(19) $$\sum \beta_i \zeta_i - \mathsf{E}(\sum \beta_i \zeta_i) = \sum (N\beta_i - C \operatorname{sign} u_i) n_i$$

and one gets easily

(20) $$\mathsf{E}(\sum \beta_i \zeta_i - \mathsf{E}(\sum \beta_i \zeta_i))^2 = N(N^2 - C^2)\sigma^2 .$$

From (16), (20)

(21) $$\frac{(\mathsf{E}(\sum \beta_i \zeta_i))^2}{\mathsf{E}(\sum \beta_i \zeta_i - \mathsf{E}(\sum \beta_i \zeta_i))^2} = \left(\frac{S}{\sigma}\right)^2 \frac{(ND - CH)^2}{N(N^2 - C^2)} .$$

Let especially (as in radar)

(22) $$\operatorname{sign} s_i = \beta_i$$

Then from (16) to (21)

(23) $$\frac{(\mathsf{E}(\sum \beta_i \zeta_i))^2}{\mathsf{E}(\sum \beta_i \zeta_i - \mathsf{E}(\sum \beta_i \zeta_i))^2} = \left(\frac{S}{\sigma}\right)^2 \frac{N^2 - C^2}{N} = \left(\frac{S}{\sigma}\right)^2 N \frac{N^2 - C^2}{N^2} .$$

Let $C = 0$ in (11), that is, the sequences in (2), (8) are not correlated. Then the second term in the middle in (23) is N, the known signal/noise gain (in the case of $U = 0$ in (2) considered also in the literature) of using the sequence (1) instead a single term. The last term on the right of (23) represents the loss in the signal/noise ratio caused by the interfering sequence.

Let $C = \pm N$, that is $\operatorname{sign} u_i = \pm \beta_i$, $(i = 1, \ldots, N)$. Then, as is to be expected, the sequence (1) cannot be detected, the expression in (23) is 0. It is interesting that the value of U itself is not contained in (21), (23), only the similarity of both sequences in (2), (8) in the sense of (11) being important. In what follows, there will be supposed

(24) $$0 < |C| < N .$$

Suppose that the sequences (8) and $\{\operatorname{sign} u_i\}$ are coincident in $N - m$ terms, the remaining m terms being of opposite signs. Then $C = N - 2m$ and the second term in the middle of (23) is

(25) $$G = \frac{4m(N - m)}{N}$$

and for $N \gg m$ and $m = 1, 2, \dots$ one gets approximately $G = 6, 9, 11, \dots, 6 + + 10 \log m$ (dB).

Let now (22) be not fulfilled. Returning to (16) one sees that (since the denominator in (21) is independent on (1)) (21) is maximum precisely if

$$\text{sign } s_i = \text{sign } (N\beta_i - C \text{ sign } u_i) \tag{26}$$

or

$$\text{sign } s_i = - \text{sign } (N\beta_i - C \text{ sign } u_i) . \tag{27}$$

There will suffice here to investigate the case of (26). Let $\beta_i = 1$. Then with respect to (24) sign $s_i = 1$. Let $\beta_i = -1$. Then sign $s_i = -1$. And vice versa. Thus the best signal detection is obtained for matched sequences (1) and (8), a well known result for $U = 0$, and a well known more-channel scheme can be used to discriminate more known sequences (1). Moreover not knowing the sequence in (2), CH on the right of (16) can be positive or negative. Thus one can argue that the best what can be done in this case is to choose the sequences (1), (8) so that $D = 0$ in (16), again a well known case of orthogonal sequences.

In what follows, only the case of the detection of a single known sequence will be investigated.

3. THE DETECTION WITH KNOWN σ

Let (22) hold. Then from (16), (17), (18), (20) and the supposition on the noise sequence (3), $\sum \beta_i \zeta_i$ is Gaussian ($\equiv \mathrm{N}(\cdot, \cdot)$)

$$\sum \beta_i \zeta_i \equiv \mathrm{N}(S(N^2 - C^2) , \quad \sigma . \sqrt{(N . (N^2 - C^2))}) \tag{28}$$

and

$$\frac{\sum \beta_i \zeta_i}{\sigma \sqrt{(N(N^2 - C^2))}} \equiv \mathrm{N}\left(\frac{S}{\sigma} \sqrt{\frac{N^2 - C^2}{N}}, 1\right). \tag{29}$$

Let firstly $S = 0$. Let p_f be the prescribed false alarm probability. Then k_1 is to be found so that

$$\Phi(k_1) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{k_1} e^{-t^2/2} \, dt = 1 - p_f . \tag{30}$$

Further, let p_d be the prescribed detection probability. Then k_2 is to be found so that

$$\Phi(k_2) = 1 - p_d . \tag{31}$$

From (29), (30), (31)

$$k_1 + (-k_2) = \frac{S}{\sigma} \sqrt{\frac{N^2 - C^2}{N}} \tag{32}$$

and the necessary S/σ can be found from this equation. Or, knowing S/σ, one can progress in reverse sense.

Actually, for signal detection $\sum \beta_i \zeta_i$ is computed from the observed values and compared with the threshold (see (28))

(33) $$\sum \beta_i \zeta_i > k_1 \, . \, \sigma \, . \, \sqrt{(N(N^2 - C^2))} \, .$$

Clearly, $\sum \beta_i \zeta_i$ is the numerator in (12) and C is given by the middle term in (11). σ is supposed known or, what is practically the same, it can be estimated independently on the sequences (1), (2), which cannot be easily done in all circumstances.

It is also seen from (28) that for s of (12)

(34) $$s \equiv \mathsf{N}\left(S, \sigma \sqrt{\frac{N}{N^2 - C^2}}\right).$$

Further, there is clear that for sufficiently great N the above formulae hold approximately also if the noise in (3) is not Gaussian, according to the central limit theorem.

4. THE DETECTION WITH UNKNOWN σ

If σ is unknown or, what is practically the same, cannot be easily measured independently on other signal components, there may be expected that some variant of "Student's t" will be valid. In what follows, it will be supposed $\sigma = 1$ without loss of generality.

If the random variables ζ_i in (14) were independent for $S = 0$, then the expression

(35) $$\sqrt{(N - 1)} \frac{\sum \beta_i \zeta_i}{\sqrt{(N \sum \zeta_i^2 - (\sum \beta_i \zeta_i)^2)}}$$

would possess the t-distribution. But the variables ζ_i are dependent so that a modification of (35) is needed.

From (15)

(36) $$\sum \beta_i \zeta_i = N \sum \beta_i n_i - C\psi \, ,$$

(37) $$\sum \zeta_i^2 = N^2 \sum n_i^2 - N\psi^2 \, ,$$

where

(38) $$\psi = \sum \operatorname{sign} u_i n_i \, .$$

Obviously

(39) $$\sum \beta_i n_i \equiv \mathsf{N}(0, \sqrt{N}) \, ,$$

(40) $$\psi \equiv \mathsf{N}(0, \sqrt{N}) \, .$$

Further,

$$\mathsf{E}(\psi \,.\, \sum \beta_i n_i) = C\,. \tag{41}$$

From (39), (40), (41), the correlation coefficient

$$\varrho(\psi, \sum \beta_i n_i) = \frac{C}{N}\,. \tag{42}$$

Since both variables in (42) are Gaussian, there can be posed

$$\psi = \frac{C}{N} \sum \beta_i n_i - \xi\,, \tag{43}$$

where ξ is a new Gaussian random variable not correlated with, thus also independent on $\sum \beta_i n_i$. From (39), (40), (43) there follows $\mathsf{E}(\xi) = 0$, further with (41)

$$\mathsf{E}(\xi^2) = \mathsf{E}\left(\frac{C}{N} \sum \beta_i n_i - \psi\right)^2 = \frac{C^2}{N^2} N + N - 2\frac{C}{N} C = N - \frac{C^2}{N}\,. \tag{44}$$

Thus ξ is a Gaussian variable

$$\xi \equiv \mathsf{N}\left(0, \sqrt{\left(N - \frac{C^2}{N}\right)}\right) \tag{45}$$

fulfilling the condition to be independent on $\sum \beta_i n_i$. From (36), (43)

$$\sum \beta_i \zeta_i = \left(N - \frac{C^2}{N}\right) \sum \beta_i n_i + C\xi\,. \tag{46}$$

Similarly from (37), (43)

$$\sum \zeta_i^2 = N^2 \sum n_i^2 - \frac{C^2}{N} (\sum \beta_i \zeta_i)^2 - N\xi^2 + 2C\xi \sum \beta_i n_i\,. \tag{47}$$

Using (46), there will be further computed

$$\frac{N^2}{N^2 - C^2} (\sum \beta_i \zeta_i)^2 = (N^2 - C^2)(\sum \beta_i n_i)^2 + \frac{N^2 C^2}{N^2 - C^2} \xi^2 + 2NC\xi \sum \beta_i n_i\,. \tag{48}$$

Now, from (47), (48)

$$N \sum \zeta_i^2 - \frac{N^2}{N^2 - C^2} (\sum \beta_i \zeta_i)^2 = N^3 \left(\left(\sum n_i^2 - \frac{1}{N} (\sum \beta_i n_i)^2\right) - \frac{N}{N^2 - C^2} \xi^2\right). \tag{49}$$

But, it is well known ([3], p. 382) that the first term in the parentheses on the right possesses the χ^2-distribution with $N - 1$ degrees of freedom and $\sum \beta_i n_i$ is independent on it. The same independence will now be imposed on ξ as a supplementary condition, without loss of generality.

Further, it follows from (45) that the second term in the parentheses on the right is distributed as χ^2 with one degree of freedom.

From (46), (49)

$$(50)\qquad \frac{\sum \beta_i \zeta_i}{\sqrt{\left(N \sum \zeta_i^2 - \frac{N^2}{N^2 - C^2} (\sum \beta_i \zeta_i)^2\right)}} =$$

$$= \frac{\frac{N^2 - C^2}{N} \sum \beta_i n_i + C\xi}{N\sqrt{N}\sqrt{\left(\sum n_i^2 - \frac{1}{N} (\sum \beta_i n_i)^2 - \frac{N}{N^2 - C^2} \xi^2\right)}}.$$

The numerator on the right has mean value 0 and its variance is from (39), (45) $N(N^2 - C^2)$. Thus standardizing it to 1 and dividing further by $\sqrt{(N - 1)}$ in the numerator, one gets

$$(51)\qquad \frac{\sum \beta_i \zeta_i}{\sqrt{((N^2 - C^2)/N^2(N - 1))}\sqrt{\left(N \sum \zeta_i^2 - \frac{N^2}{N^2 - C^2} (\sum \beta_i \zeta_i)^2\right)}} =$$

$$= \frac{X}{\sqrt{\dfrac{\left(\sum n_i^2 - \frac{1}{N} (\sum \beta_i n_i)^2 - \frac{N}{N^2 - C^2} \xi^2\right)}{N - 1}}}.$$

Here, $X \equiv \mathsf{N}(0, 1)$. For $N \gg 1$, one may neglect the term containing ξ^2 under the root sign on the right. Then ([3], p. 237 and 387), the variable on the right possesses approximately the t-distribution with $N - 1$ degrees of freedom. For smaller N, correcting slightly the influence of the neglected term by the factor $(N - 2)/(N - 1)$ respecting the mean value, one gets finally the detection rule as follows:

the signal is declared as present if

$$(52)\qquad \sum \beta_i \zeta_i > t_{N-1}(\alpha) \sqrt{\left(\frac{N^2 - C^2}{N(N - 2)} \sum \zeta_i^2 - \frac{1}{N - 2} (\sum \beta_i \zeta_i)^2\right)},$$

where t is the tabulated value with $N - 1$ degrees of freedom and $\alpha = p_f$. Exceptionally, $N = C$ caused by noise and this case shall be eliminated in advance in programming (52).

Now, if the sequence (1) is present, then one sees from (20) that the variance remains unchanged and only the mean is changing. By an easy calculation one finds that the auxiliary variable X in (51) possesses now the expected value (with $\sigma \neq 1$ again introduced)

$$(53)\qquad \mathsf{E}(X) = \frac{S}{\sigma} \sqrt{\frac{N^2 - C^2}{N}}$$

(cf. with (32)) and the variance 1.

Thus the standard tables [4], [5] of the noncentral t can be used and with the tables notation

$$(54)\quad \frac{S}{\sigma}\sqrt{\frac{N^2 - C^2}{N}} = \delta(N-1, t_0, \varepsilon) = t_0 - \lambda(N-1, t_0, \varepsilon)\sqrt{\left(1 + \frac{t_0^2}{2(N-1)}\right)},$$

where $t_0 = t_{N-1}(\alpha) = t_{N-1}(p_f)$ and $\varepsilon = p_d$. For $\varepsilon > 0{\cdot}5$, as is usually the case,

$$(55)\qquad \delta(N-1, t_0, \varepsilon) = -\delta(N-1, -t_0, 1-\varepsilon).$$

Conversely, given S/σ and t_0, one can find ε.

Given p_f, p_d and computing k_1, k_2 from (30), (31) and δ from (54),

$$(56)\qquad l = \left(\frac{k_1 - k_2}{\delta}\right)^2 < 1$$

represents the loss of signal/noise ratio caused by not knowing σ.

5. SOME RESULTS OF SIMULATION

To show the applicability of previous results for small N a short extract of simulation results will be given in following two tables. In both cases, $N = 7$ has been chosen. The sequence (1) is the known Barker-7 sequence. There are two interfering sequences (2):

$$(57)\qquad \begin{aligned} &\text{a)}\ \{4, 4, 4, -4, -4, -4, -4\},\\ &\text{b)}\ \{4, 4, -4, -4, 4, 4, 4\}. \end{aligned}$$

The Gaussian sequence (3) has been approximated as the sum of 5 pseudorandom numbers with uniform distribution in the interval $\langle -1, +1 \rangle$. Since the standard deviation of this sum is $\sqrt{(5/3)}$, the sum of pseudorandom variables has been multiplied by such a factor, that the desired s/n ratio (with Barker-7 as the wanted sequence) be obtained. In both tables $p_f = 0{\cdot}05$ and $p_d = 0{\cdot}7$. For the table 1 (known σ) one gets from (30) $k_1 = 1{\cdot}64$. From (31), $k_2 = -0{\cdot}52$. Since $S = 1$ and $C = 5$ for (57) a) and $C = 1$ for (57) b), one obtains in (33) $k_1\sigma = 1{\cdot}41$ in the case a) and $1{\cdot}97$ in the case b). From (32), one gets factors to multiply the sum of pseudorandom numbers as resp. $0{\cdot}66$ and $0{\cdot}94$.

For the case of the Table 2 (unknown σ), one finds in the tables of "Student's t" to $p_f = 0{\cdot}05$ the one-sided value $t_6(0{\cdot}05) = 1{\cdot}94$. From (54), (55), one gets the factors to multiply the sum of pseudorandom numbers as resp. $0{\cdot}59$ and $0{\cdot}83$. And $\delta = 2{\cdot}45$ from the tables of the noncentral t.

The agreement of computed and simulated results (on the basis of simple $np \pm \pm 2\sqrt{(npq)}$ criterion) is good.

Table 1. (3 × 100 simulations).

$k_1\sigma$	$100p_f$	$100p_d$	1. interference		2. interference	
			without sig.	with sig.	without sig.	with sig.
1·41 (1. interf.)	5	70	4	66	6	68
1·97 (2. interf.)			4	71	4	76
			5	73	6	61

Table 2. (3 × 100 simulations.)

$t_6(0{\cdot}05)$	$100p_f$	$100p_d$	1. interference		2. interference	
			without sig.	with sig.	without sig.	with sig.
1·94	5	70	3	66	5	70
			6	70	6	71
			2	70	8	67

There is seen that to get the same detection probability 0·7, s/n ratio is 3 dB greater for (57) a) than for (57) b). From (56), the loss of s/n due to not knowing σ is about 1 dB.

6. CONCLUDING REMARKS

If σ can be computed independently, then computing (33) in real time is substantially less laborious than (52), as is immediately seen.

The results of this article can be generalized in several ways, e.g. the random variables s, u from (12), (13) can be used with an appropriate threshold to decide whether the present method or the usual matched filter is to be used for the detection.

ACKNOWLEDGEMENT

For all computer programming, computings and simulations, and discussions of the results, the author is indebted to Ing. J. Beneš, CSc.

(Received June 1, 1981.)

REFERENCES

[1] J. L. Lewis: Bit Error Probability of a Large-Jammer Cancellation System. NTC 78 Birmingham Conf. Rec., IEEE, 1978, V. 3, 35.3.1—35.3.5.

[2] V. P. Afanasjev: Digital algorithms of discrimination of phase-manipulated wide-band signals (in Russian). Izv. VUZov SSSR — Radioelektronika *22* (1979), 8,65—72.

[3] H. Cramér: Mathematical Methods of Statistics. Princeton Univ. Press, Princeton, N.J 1946.
[4] N. L. Johnson, B. L. Welch: Some applications of the noncentral *t*-distribution. Biometrika *31* (1939), 362—389.
[5] J. Janko: Statistical Tables (in Czech). NČSAV, Prague 1958.

Dr. Ludvík Prouza, DrSc., Tesla — Ústav pro výzkum radiotechniky (Institute of Radioengineering), Opočínek, 533 31 p. Lány na Důlku. Czechoslovakia.