DETECTION & ESTIMATION OF COVERT DS/SS SIGNALS USING HIGHER ORDER STATISTICAL PROCESSING

Mamdouh Gouda, Ernest R Adams,& Peter C J Hill

School of Engineering & Applied Science (SEAS)
Cranfield University RMCS, Shrivenham, Swindon, Wilts, SN6 8LA, UK
Phone: +44(0)1793 785211 Fax: +44(0)1793 783192
E-mail: gouda@rmcs.cranfield.ac.uk

Abstract

Conventional linear and non-linear receivers are generally ineffective in detecting direct-sequence spread spectrum (DS/SS) signals if the spreading sequences are unavailable. An investigation into using correlation-based processing is reported showing that the cyclostationary property of DS/SS provides detection capability. Finally we describe with results an emerging technique based on higher-order statistics where triple correlation analysis is used, leading to the detection and estimation of DS/SS length and its code generating function *g(X)*.

1. Introduction

Modern civil and military electronic communication systems are becoming ever smarter, more complex and in some cases also very difficult to intercept because the nature of the design leads to a covert signal structure. In fact some transmitted waveforms are intentionally designed to make the detection process virtually impossible. Such low-probabilities-of-intercept (LPI) signals have very wideband extremely low power spectral density signatures based on a hidden code structure; moreover they can operate in a complex radio environment of high noise, interference, jamming and other co-channel signals.

A particularly difficult threat signal to intercept is direct-sequence spread spectrum (DS/SS) and in this case the problem is exacerbated when such systems operate in multiple access mode using code division multiple access (CDMA). Moreover, if the pseudo-noise spreading codes are very long and the intercept window is short including an unknown number of aperiodic data modulations then standard signal processing methods based on second order statistics are severely limited.

State of the art techniques for communications signal detection which are based on second-order (i.e. variance) spectral processing are considerably limited by 'phase blindness' and the inability to easily separate out wanted signals from background noise[1]. Our study, however, is focused on higher-order statistical processing which specifically uses the cyclostationary signal property but is combined with the suppression of Gaussianity [2,3] in order to improve the SNR of the detection process. This was the key motivation for our attack on the problem of detecting DS/SS using correlation analysis and also spectral processing for detecting CDMA and chip-code characterisation respectively. The paper discusses the theory of triple correlation function analysis as applied to the detection of DS/SS PN chip code sequences in some detail and then describes the various methods which have been investigated for estimating the basis code polynomials of DS/SS signals in the presence of channel noise. Attention is given to the importance of the *doubling* technique which improves SNR and reduces the dimensional@ of tcf characterization.

2. Higher-order moment signal processing techniques

The proposed detector for covert DS/SS signals uses third order cumulants in the form of triple correlation analysis, bispectral processing and an associated characterisation process[4]. HOM/HOS techniques should be better able to exploit the non-gaussian cyclostationarity of the signal against the channel noise and interference. For example, the shift-and-add property of the m-sequences[5], i.e. $u \oplus u_p = u_q$, where $\oplus$ is binary addition of sequence elements, with an equivalent polynomial representation,

$$g(X) + X^p g(X) \mathrm{mod}(X^L + 1) = X^q g(X) \mathrm{mod}(X^L + 1),$$

leads to the delta function response for the periodic autocorrelation function (ACF) $C_{xx}(\tau) = \mathrm{E}[v(t)v(t+\tau)]$ and this ACF is no more than the second order cumulant in HOM terms.

Higher-order cumulants simply extend the averaging process by considering additional time-shifted versions of the same m-sequence signal. In particular, the third-order cumulant or *triple correlation* is defined as

$$C_{xxx}(\tau_1, \tau_2) = \mathrm{E}[v(t)v(t+\tau_1)v(t+\tau_2)]$$

where $\tau_1 = pT_c$ and $\tau_2 = qT_c$ for C_{xxx} sampled at the chip rate $1/T_c$ Hz. In practice, m-sequences use the values $\pm$ 1 rather than 0 and 1: $(0,1) \rightarrow (1,-1)$. The previously defined binary addition of sequences, $\oplus$, is equivalent to multiplication, $*$, in this new domain.

3. Triple correlation of complete m-sequences

The discrete version of $C_{xxx}(\tau_1, \tau_2)$ is evaluated as

$$C(p,q) = \frac{1}{L}\sum_{i=1}^{L} v(i)v_p(i)v_q(i)$$

where $v_j(i) = 1$ if $u_j(i) = 0$ and -1 if $u_j(i) = 1$, i denoting the i-th element of the sequence. By the shift and add property, for certain (p', q'), $u_{p'} \oplus u_{q'} = u$, and it follows that $\forall i$, $v_{p'}(i)v_{q'}(i) = v(i)$. For those pairs:

$$C(p',q') = \frac{1}{L}\sum_{i=1}^{L}[v(i)]^2 = 1.$$

For other (p, q) pairs, $v_p * v_q = v_s$, where $v_s \neq v$ but is an m-sequence by field closure, in which case

$$C(p,q) = \frac{1}{L}\sum_{i=1}^{L} v(i)v_s(i) = -1/L.$$

Thus the shift and add property results in $C(p,q)$ peaks at locations for which $\alpha^{p'} + \alpha^{q'} = 1$. Each peak at (p', q') is mirrored at (q', p'), as $\alpha^{q'} + \alpha^{p'} = 1$. Also, because of the existence and uniqueness of a q'for eachp'in the range $1 \leq p', q' \leq$ L-l, there is exactly one peak in each row and column between 1 and L-1, a total of L-l peaks.

In the original equation for $C(p,q)$, v,(i) = $v(i+p)$ and $v_q(i)=v(i+q)$, i.e. those sequences are advanced in time or shifted to the left (LS). There is a corresponding delayed or right-shifted (RS) version of

C. Substituting $i+q=n$ and assuming $L>q\geq p\geq 0$, $v(i+q)=v(n)$, $v(i+p)=v(n-(q-p))$ *and* $v(i)=v(n-q)$. Thus $C(p,q)$ may be written in terms of RS versions of v:

$$C(p,q)=\frac{1}{L}\sum_{n=1}^{L}v(n)v_{-q}(n)v_{-(q-p)}(n) \qquad \text{[RS]}$$

$$=\frac{1}{L}\sum_{n=1}^{L}v(n)v_{L-q}(n)v_{L-(q-p)}(n) \qquad \text{[LS]}$$

For each peak at (p,q) there is a corresponding peak at $(L-q,\ L-(q-p))$; reflections are also peaks. E.g. as sequence *a* ,with generating polynomial $g(X)=X^5+X^2+1$, (L=31) has peak (p,q) =(1,18), it also has peak (13,14); (18,1) and (14,13) are also peaks as shown in Fig 1.

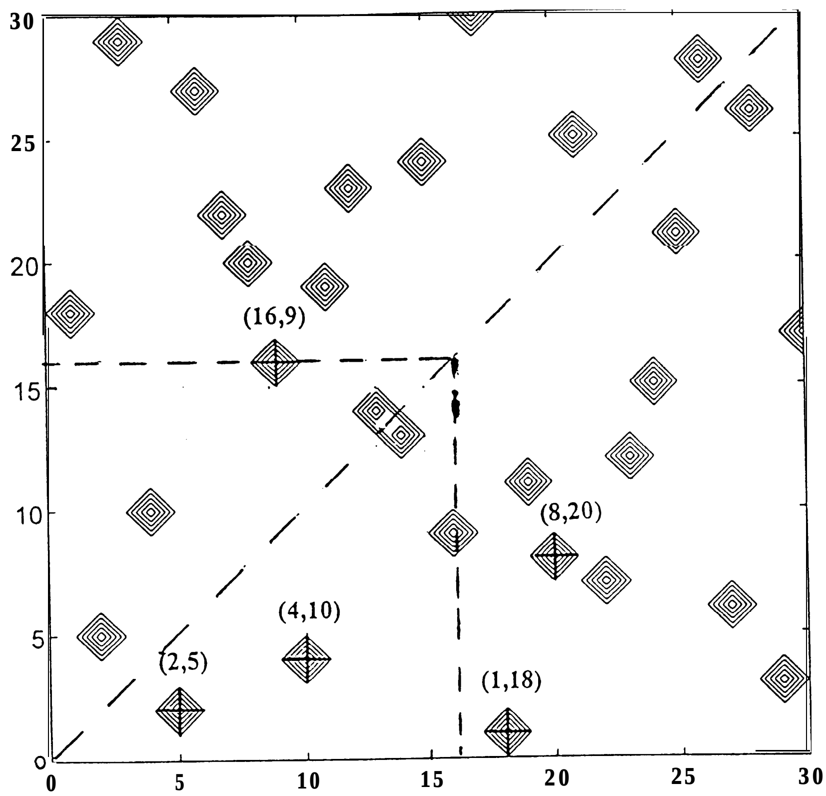

Fig 1. Coset summing for partial tcf [45]

4. Determination of *g(X)* from triple correlation peak locations

The theory developed so far assumes the m-sequence length L is known. In particular, knowledge of L is necessary to evaluate the triple correlation C(p, q). However, there is evidence that sufficiently long partial m-sequences produce good estimation of peak locations. L may be derived from the peak locations

$$(i, j), (2i, 2j), \ldots, (2^k i, r) \quad \text{for } i < j$$

Assuming the last pair is the first for which $r < 2^k j$, $r = 2^k j \bmod(L)$, or $L = 2^k j - r$. As an example, the pairs (1, 18),(2, 5) would produce L = 36 - 5 = 31. As peaks may be displaced, several such seqıences should be examined.

If L is known, it is possible to determine $g(X)$, and thus the tap weights of its LFSR, from a *single* triple correlation peak location. This follows from the fact that, for a given L, different primitive $g(X)$ produce no common peaks.

As an illustration, consider the m-sequences of length 31 generated by $g(X) = X^5 + X^2 + 1$ (45 octal). If, for example,as shown in Fig 1, the peak location (1, 18) is known, the following peak locations may be predicted:

$$(1, 18), (2, 5), (4, 10), (8, 20), (16, 9).$$

From the one peak such as (2, 5) we can derive $g(X)$:

$$\left.\alpha^2 + \alpha^5 = 1 \Rightarrow \alpha^5 = \alpha^2 + 1\right\} \Rightarrow \alpha^5 + \alpha^2 + 1 = 0.$$

Thus a is a solution of $g(X) = 0$, where $g(X)$ must be of order 5 and include 1:

$$g(X) = X^5 + X^2 + 1.$$

5. Coset summing for better detection

Although the effects of noise may be reduced by averaging the tcfs from several short signal samples, this process is computationally costly and assumes persisting m-sequences. Coset summing is an alternative or complementary technique which improves SNR and reduces the dimensionality of tcf characterization. It may lead either to powerful multivariate discrimination of fragments of known m-sequences or to the blind determination of an m-sequence from the detection of a single tcf peak.

Coset summing involves searching NxN partial tcfs for peaks by using the *doubling* property of peak locations. Each feasible tcf location is visited, beginning (p,q)=(1,2),(1,3),...,(2,3),(2,4),..., and the following sum of non-repeated tcf values calculated:

$$S = C'(p,q) + C'(2p \bmod L, 2q \bmod L) + \ldots + C'(2^r p \bmod L, 2^r q \bmod L)$$

where $p=2^{r+1}p \bmod L$ and $q=2^{r+1}q \bmod L$. Each doubled location, $(2^i p \bmod L, 2^i q \bmod L)$ for $1 \leq i \leq r$, is excluded from the future search as its coset is represented by the initial location (p,q). Thus each coset present in the partial tcf is represented by a single peak, called the coset leader. Clearly, all the peaks of a coset will not generally be present in the partial tcf. When doubled locations lie outside the partial tcf $(p', q' > N)$ no contribution is made to the sum but doubling is continued, and further values for locations within the partial window summed, until the original (p,q) results.

Coset summing is illustrated in Fig 1, the tcf of a 3 1 length m-sequence [45]. Assuming only a 16x16 partial window is available, the peak at (2,5) would be the first peak encountered in the search. The other coset peaks added would be at (4,10) and (16,9): doubling would generate the other coset members (8,20) & (1 ,18) lying outside the partial window. However, the peaks at (4,10) & (16,9) would be excluded from the search and their contributions included at (2,5), the coset leader.

Clearly, in addition to reducing dimensionality, the use of coset sums of available tcf values will improve the detectability of actual peaks. Actual peaks of partial tcfs of m-sequences in noise have average values of 1 while the average non-peak values are $-1/L$. If the search begins with a non-peak location, other locations generated by doubling will also be non-peak.. Thus all the coset sums will consist of exclusively peak values or exclusively non-peak values. These summed values may be tested against a threshold to decide whether they arise from a coset of actual peaks.

6. Conclusions

Triple correlation analysis provides a powerful means of searching for and detecting the presence of covert wideband signals such as DS/SS. The results show that the tcf is an excellent basis for detection and identification of m-sequences. The doubling process (coset sum)improves the detectability of actual triple correlation peak and reduce non-peak values: However higher-order statistical processing can extract more information than that conveyed by second-order power spectral density or autocorrelation function.

7. References

1. Stephens J.P., 'Advances in signal processing technology for electronic warfare', Journal of Electronic Defense, September 1995.

2. Gardner W.A., 'Signal interception: a unifying theoretical framework for feature detection', IEEE Trans on Comms, 36, 1988.

3. Nikias C. L. & Petropulu A. P., 'Higher-order spectra analysis - a nonlinear signal processing framework', PTR Prentice-Hall Inc., 1993.

4. E R Adams, M Gouda & P C J Hill, 'Detection and characterisation of DS/SS signals using higher-order correlation', Proceedings of IEEE 4th ISSSTA, Mainz, Germany,Sep 1996, pp27-3 1.

5. S. W.Golomb, Shift Register Sequences, Rev.Ed.(Aegean Park Press, 1982).