DETECTION OF DIRECT SEQUENCE SPREAD SPECTRUM SIGNALS USING HIGHER-ORDER STATISTICAL PROCESSING

Ernest R Adams & Peter C J Hill

Cranfield University RMCS, Shrivenham, Swindon, Wilts, SN6 8LA UK

ABSTRACT

Direct-sequence spread spectrum signals are undetectable by conventional intercept receivers. However, the properties of the m-sequences used to generate such signals give rise to characteristic patterns of triple correlation peaks. These patterns remain discernable for short signal intercepts in the presence of data modulation, other m-sequences and noise, allowing the detection of individual m-sequences.

1. INTRODUCTION

Interception of covert signals such as direct sequence spread spectrum (DS/SS) is becoming increasingly vital for the radio monitoring and surveillance authorities as both legitimate and criminal communities start to employ these techniques. It is often important to estimate the code structure of such transmissions. Conventional linear and (second order) nonlinear search and detection receiver processors have very limited performance in this respect [1]. Partial code sequences with large spreading ratios present special problems in detection. Our approach is based on the use of higher-order statistical (HOS) analysis in the time domain concentrating on triple correlation which is robust against additive background noise [2]. The initial investigation focuses on m-sequence DS/SS where the discrete polynomial code structure is theoretically available for validation of the computer simulation results. The techniques described allow detection and chip code structure (generator) estimation of full and partial m-sequences, including the effects of data switching, in additive noise channels; multiple signals in code division multiple access (CDMA) are also considered.

This work is supported by US Air Force, Wright Laboratory, Contract F61708 95 C008.

2. TRIPLE CORRELATION OF DS/SS SIGNALS

It is known that the triple correlation function (tcf) is useful for distinguishing m-sequences: their shift and add property produces tcf peaks. The locations of tcf peaks are used to find the linear shift register feedback function which generates an m-sequence in [3], which also briefly considers the presence of noise and data modulation. Galois field theory may be used to derive the generator polynomial of an m-sequence from a single exact tcf peak location. As an exact peak location will not be known in practice, a detailed analysis is given for the tcf's of partial m-sequences and the presence of additive noise.

It may be shown that the tcf of two superimposed m-sequences consists of the two individual sets of peaks of height 1 with additional peaks depending on cross-correlation properties of the sequences [4]. For sequences of length L=31, these cross-correlation peaks have magnitudes between 1/31 and 11/31. Figure 1 shows the tcf of two superimposed preferred sequences: a threshold ensures just the two sets of individual peaks are present, albeit modified in amplitude by cross-correlation terms. The theory may easily be extended to more than two sequences.

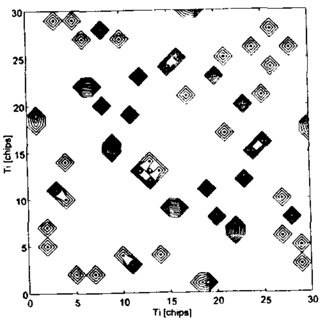

Figure 1: Contour tcf of the superposition of two preferred m-sequences of length 31.

In practice, data sample lengths will be shorter than the unknown lengths of any complete sequences present. If only a partial sequence of length $N(<L)$ is sampled, perdiodicity may not be assumed in calculating its triple correlation. A partial tcf may be estimated:

$$C'(p,q)=\frac{1}{N-q}\sum_{i=1}^{N-q} v(i)v(i+p)v(i+q)$$

for $0 \le p \le q \le N$-1

$$C'(q,p)=C'(p,q) \quad \text{for } 0 \le q < p \le N\text{-}1$$

If the true $C(p,q)$ has a peak at (p',q') then $\forall i$, $v(i)v(i+p')v(i+q')=1 \Rightarrow C'(p',q')=1$, ie $C'(p,q)$ has the same peaks as $C(p,q)$ for $p,q \le N$-1. For other (p,q) pairs, $v(i+p)v(i+q)=v(i+r)$ for some other shift $r(\neq 0)$, so

$$C'(p,q)=\frac{1}{N-q}\sum_{i=1}^{N-q} v(i)v(i+r)=C'(r)$$

The partial correlation coefficient $C'(r)$ based on an average of N-q products for an L-length m-sequence has a mean value of -1/L and a variance of approximately 1/(N-q) for N-q<<L, ie the variance is inversely proportional to the sequence length used. Thus the quality of estimates depends on the absolute sample length and not on the proportion of the sequence covered by the sample. Figures 2 and 3 show excellent agreement because the partial tcf in figure 3 is based on a long partial sequence of length 256, even though this is a small sample of the complete length of 1023 used to calculate the tcf in figure 2.

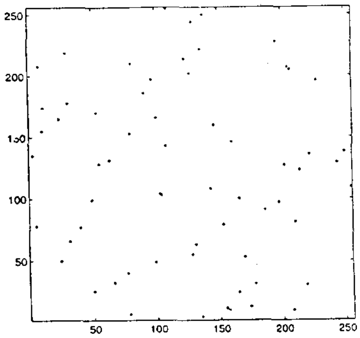

Figure 2: Part of the tcf of the complete 1023-length m-sequence generated by $X^{10}+X^8+X^3+X^2+1$.

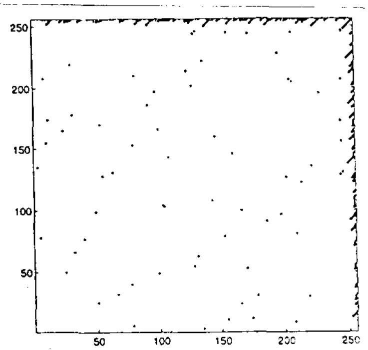

Figure 3: Partial tcf for 256/1023 values of the sequence in figure 2 (threshold = 0.7).

The noise theory developed shows that the tcf of an m-sequence plus Gaussian noise has the same expected value as the m-sequence tcf except for a small bias on its diagonal. Variances depend on the noise variance, with diagonal values three times greater than off-diagonal. The tcf of two or more m-sequences in Gaussian noise is more confused by extra spurious peaks due to additional bias terms and increased variances. Figure 4 shows the effect of adding noise to the partial sequence used in figure 3. Detectability of peaks may be enhanced by averaging tcf's, reducing noise variances and reinforcing principal peaks due to persisting sequences. The following detection method uses multivariate statistical measures combining multiple peak locations.

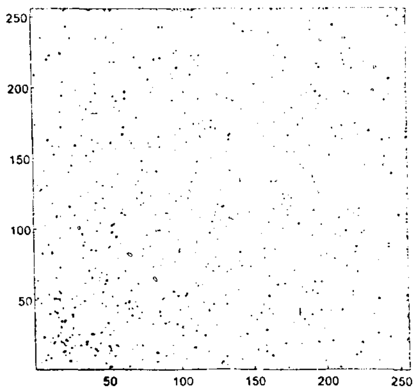

Figure 4: Partial tcf for the sequence in figure 3 + noise, SNR 3 dB (threshold = 0).

3. STATISTICAL DETECTION OF DS/SS SIGNALS

Signals to be detected are assumed to be represented by partial samples from m-sequences of length 511 or 1023. Partial samples are considered with and without data modulation. However, short signal intercepts are unlikely to contain more than one data transition, beyond which all values are assumed to be changed in sign. Without synchronization information, the data transition must be considered to occur at an arbitrary position in the m-sequence sample. Figure 5 shows the effects of data modulation on the partial tcf of the sequence used in figure 3. The modulation occurs after 128 of the 256 signal samples (from 1023). Comparing with figure 2, all peaks close to the diagonal (*p*=*q*) are preserved for shifts >100. Even for a small fraction of the sequence and data modulation in the worst place, much of the tcf peak information is preserved. The problem of detection in the presence of an unknown modulation remains, given the modulation's influence on the tcf peaks.

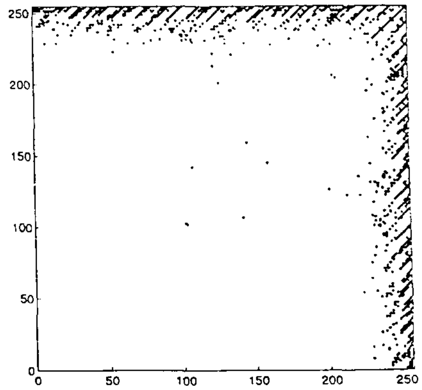

Figure 5: Partial tcf for the sequence in figure 3 data-modulated after 128 values (threshold = 0.4).

Without data modulation, detecting sequences is possible in noise levels worse than 0 dB. This is illustrated in figure 6, showing the clear separation of samples from two partial (128/511) m-sequences with 0 dB noise. The two discriminant directions are sums of tcf values at known peak locations for the two m-sequences respectively. The discriminant space may be extended to cover all possible sequences present. Figure 7 shows the poorer separation at 0 dB when data modulation occurs after 128 chips of two 256-length partial (256/1023) m-sequences. The discriminant directions were calculated using multivariate analysis of variance [5], separating variances/covariance contributions due to the effects of sequence, modulation position, their interaction etc. The best overall discriminant directions in the presence of unknown modulation position are linear combinations of tcf values close to the diagonal, as anticipated.

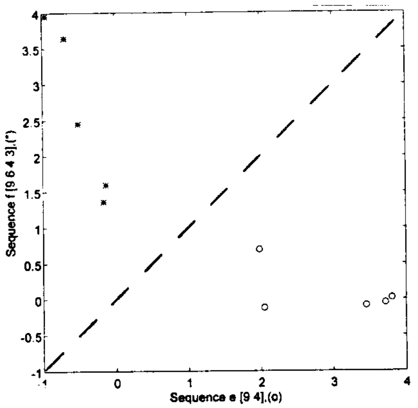

Figure 6: Discrimination of partial sequences e & f (128/511), SNR 0 dB.

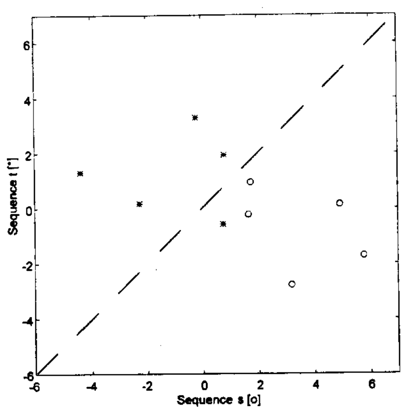

Figure 7: Discrimination of partial sequences s & t (256/1023) data-modulated after 128 values, SNR 0 dB.

4. CONCLUSIONS

Detection of single and also multiple DS/SS signals using tcf processing has been developed taking account of short intercept windows, channel noise and data modulation. Simple discriminant analysis based on sums of tcf values at known peak locations is shown to be effective at low SNRs. A more sophisticated statistical analysis produces modulation independent detection but with reduced performance.

5. REFERENCES

[1] D.L. Nicholson, “Spread spectrum signal design - LPE & AJ Systems”, Computer Science Press Inc., 1988.

[2] C.L. Nikias & A.P. Petropula, “Higher-order spectra analysis - a nonlinear signal processing framework”, PTR Prentice-Hall Inc., 1993.

[3] E.S. Warner, B. Mulgrew & P.M. Grant, “Triple correlation analysis of binary sequences for codeword detection”, IEE Proc. - Vis Im Sig Proc., 141, 5 October 1994.

[4] E.R. Adams, M. Gouda & P.C.J. Hill, “Detection and characterisation of DS/SS signals using higher-order correlation”, Proc. IEEE ISSSTA ‘96, 1, 27-31.

[5] A.M. Kshirsagar, “Multivariate analysis”, Marcel Dekker, 1978.