Class of binary cipher sequences with best possible autocorrelation function

L.C. Quynh and S. Prasad, B.Tech., M.Tech., Ph.D., Mem.I.E.E.E.

Indexing terms: *Signal processing, Codes and decoding*

Abstract: A new class of binary sequences called here '*m*-like cipher sequences' are introduced which have rather interesting properties for use as cipher sequences. These are derived by interleaving the so-called 'elementary *m*-sequences' along with some null sequences in a particular order. The resulting composite sequences are not *m*-sequences, but require deciphering and storage of a much larger number of bits for their complete prediction from the observed cipher text than that needed for predicting *m*-sequences of the same length. At the same time, the method of construction of these sequences ensures that their autocorrelation function is identical to that of an *m*-sequence of the same length.

1 Introduction

In many applications such as communications, radar and cryptography, it is required to design binary sequences with the following features: (i) have a peaky autocorrelation function, (ii) have an evenly balanced number of ones and zeros in the sequence length (or period), (iii) have a maximum 'period' or 'length' for a given shift register length, and (iv) be difficult to predict from partial observation.

The well known Barker codes yield the best possible autocorrelation function (ACF) for a given length of the sequence. Unfortunately, however, these codes are known only up to a length of 13. A well known class of binary sequences which satisfy most of the above features are the '*m*-sequences' generated by a linear feedback shift register (LFSR) using 'primitive' generating polynomials. The normalised ACF of an *m*-sequence is given by

$$\theta_m(l) = \begin{cases} 1, & l = 0 \bmod L \\ -\dfrac{1}{L}, & l \neq 0 \bmod L \end{cases}$$

where $L(= 2^m - 1)$ is the sequence length (or period) and m is the degree of the generating polynomial. The number of ones and zeros in the *m*-sequence is given by

$$L_1 = (L + 1)/2$$

$$L_0 = (L - 1)/2$$

For L sufficiently large, $L_1 \simeq L_0$. In spite of these useful properties, however, these *m*-sequences have an important drawback, in that they are not very suitable for use in encryption applications. This is because they do not satisfy requirement (iv) as discussed above. It is well known [1], in fact, that only $2m$ consecutive bits are needed to determine the initial state of the register and the primitive polynomial associated with the LFSR.

There has therefore been a great deal of interest in finding new, possibly nonlinearly constructed, shift register sequences having all the features discussed above. A nonlinear set of sequences which meet the unpredictability requirement (iv) quite well had been introduced in Reference 1. However, the behaviour of these sequences with respect to other properties is not well understood. More recently, Jennings [2] has introduced a class of multiplexed sequences which are claimed to satisfy some of these requirements more effectively. However, the ACFs of these sequences have sidelobes larger than those of *m*-sequences. The same is also true of a class of nonlinear sequences called the 'bent function sequences', introduced in Reference 3.

In this paper we propose a new class of binary sequences which satisfy all these four requirements rather effectively. The new sequences thus have an ACF identical to that of an *m*-sequence of the same length and, in addition, have a much greater complexity (and therefore unpredictability) associated with them. We will refer to these as '*m*-like cipher sequences' in this paper. A brief outline of the paper is now given.

In Section 2 we first briefly introduce some useful preliminary concepts for the representation of *m*-sequences via an integer mapping and their decomposition into simpler (or elementary) *m*-sequences. A simple method for the calculation of the ACF of these composite *m*-sequences using the 'integer series' representation is then outlined to highlight the structure of a class of binary sequences having an ACF identical to that of an *m*-sequence. This insight is exploited in Section 3 to construct *m*-like cipher sequences. Finally, the complexity and predictability aspects of these sequences are taken up in Section 4.

2 Preliminaries

2.1 Integer series representation of m-sequences and sequences derived from them

It is well known that all the *m*-sequences generated by a primitive polynomial for a given LFSR are shifted versions of each other. Let $W(d)$ denote a reference *m*-sequence in the d-domain,* with respect to which the phase of all other shifted *m*-sequences is measured. Thus we have

$$T^r W(d): r\text{th cyclic shift of } W(d) \quad (1)$$

Therefore each cyclic shift of the reference *m*-sequence can be specified in terms of an index number r.

We now introduce the following integer series representation of an *m*-sequence and its shifted versions:

$$W(d): (0, 1, 2, 3, \ldots, L - 1) \quad (2a)$$

$$T^1 W(d) : (1, 2, 3, 4, \ldots, L - 1, 0) \quad (2b)$$

$$\vdots$$

$$T^r W(d) : (r, r + 1, \ldots, L - 1, 0, \ldots, r - 1) \quad (2c)$$

Here the L-tuples on the right-hand side are used to represent the bit values in one period of the *m*-sequence in the following sense: an integer j in the kth position is interpreted to denote that the kth bit value i_k in the first (or

Paper 4119F (E8), first received 21st November 1984 and in revised form 29th May 1985

The authors are with the Department of Electrical Engineering, Indian Institute of Technology, Delhi, New Delhi 110016, India; Mr. Quynh is on leave from PTT Vietnam on a Ministry of Education (Government of India) scholarship, and Dr. Prasad is currently on leave with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802, USA

* It would be appropriate to choose $W(d)$ as the so-called 'phase-normalised' or 'characteristic' *m*-sequence which has the in-phase decimation property $w_n = w_{2n}$ [4] (see example in Table 1).

reference) segment of the sequence is given by the zeroth or the starting bit value of the sequence $T^jW(d)$. Similarly, the kth bit lying in the next segment (i.e. $(k + L)$th bit value in the above examples) is the next bit of $T^jW(d)$ etc. In other words, the integers in this representation show the order of interleaving of the cyclic shifts of $W(d)$ to be used in the construction of the sequence.

Thus, for the above example, we have

$$i_{k+v \cdot L} \text{ for } W(d) = v\text{th bit value in } T^kW(d) \quad (3a)$$

$$i_{k+v \cdot L} \text{ for } T^rW(d) = v\text{th bit value in } T^{k+r}W(d) \quad (3b)$$

and so on, or briefly

$$i_{[k]} \text{ for } W(d) = \text{successive bit values in } T^kW(d) \quad (3c)$$

where $[\cdot]$ denotes an integer modulo L. Usually, representation of a single segment of the sequence suffices as in eqn. 2, although it is possible to represent a complete period (or length) of the sequence through an appropriately extended integer series, as follows:

$$W(d): \{(0, 1, 2, \ldots, L-1), (1, 2, \ldots, L-1, 0), \ldots\} \quad (3d)$$

This integer series representation of an m-sequence can be obviously extended to represent any class of binary sequences which are derived from m-sequences. The advantage of this representation lies in the insight it yields into the underlying structure of a given composite binary sequence derived from an m-sequence, which is not available in the direct binary representation. This mapping has earlier been shown to be very effective in studying the ACF and CCF properties of 'composite sequences' like the 'interleaved' or 'multiplexed' m-sequences etc. [5]. In the following we now introduce the composite nature of some long m-sequences in terms of some elementary m-sequences using this representation.

2.2 *Decomposition of long m-sequences into elementary m-sequences: composite m-sequences*

When m is nonprime, we can write

$$m = m_1 m_2 \quad (4)$$

and

$$L = 2^{m_1 m_2} - 1 \quad (5)$$

where L is the 'length' or 'period' of the m-sequence. It is obvious that $(2^{m_1} - 1)$ divides L [6]. Furthermore, it has been shown in Reference 7 that an m-sequence $u_L(d)$ of length L can be obtained by multiplexing some smaller (m_1-sequences) $w_N(d)$ of length N, where $N = (2^{m_1} - 1)$. Let

$$L = N \cdot S \quad (6)$$

It seems then that a total of S $w_N(d)$ are needed to yield an interleaved sequence $u_L(d)$. *Note, however, that, in order for $u_L(d)$ to be an m-sequence, not all the S elementary subsequences can be permitted to be m_1-sequences.* This is because in that event the number of ones and zeros in $u_L(d)$ will no longer be balanced. Thus, in that case, we would have

$$L_1 = \text{number of ones in } u_L(d) = \left(\frac{N+1}{2}\right)S \quad (7)$$

$$L_0 = \text{number of zeros in } u_L(d) = \left(\frac{N-1}{2}\right)S \quad (8)$$

and

$$L_1 - L_0 = S \quad (9)$$

However, for $u_L(d)$ to be an m-sequence, we must have

$$L_1 - L_0 = 1 \quad (10)$$

thus proving the point in italics above.

Consider now a sequence $u_L(d)$ obtained by multiplexing P m_1-sequences $w_N(d)$ and $(S - P)$ null sequences $0_N(d)$ of length N, where P is chosen as

$$P = \frac{(L+1)}{(N+1)} < S \quad (11)$$

It is easy to verify that for this case the required property

$$L_1 - L_0 = 1 \quad (12)$$

is indeed satisfied, so that the resulting sequence $u_L(d)$ is a possible candidate for being an m-sequence. A method of decomposition of the sequence $u_L(d)$ in terms of a set of smaller m_1-sequences $w_N(d)$ is outlined in Reference 7.

Thus, an m-sequence with a nonprime value of m can be obtained by P elementary m_1-sequences and $(S - P)$ 'null' sequences, each of length $N = (2^{m_1} - 1)$. It is convenient to represent the resulting sequence $u_L(d)$ by the following notation:

$$u_L(d) \triangleq I_P\{T^{j_1}w, T^{j_2}w, \ldots, T^{j_P}w \text{ and } (S-P) \text{ null sequences of length } N\} \quad (13)$$

where I_P denotes a Pth-order 'interleaving' or multiplexing of the indicated sequences, and where each j_i is a specifically selected value between 0 and $(N - 1)$, as indicated in the construction procedure of Reference 7. Alternatively, a more explicit and useful representation for $u_L(d)$ can be obtained by using the integer series representation discussed above, in view of the fact that $u_L(d)$ is essentially derived from (smaller) m_1-sequences. This is best illustrated by an example.

Consider the m-sequence $u_L(d)$ generated by the polynomial

$$\phi(d) = 1 + d + d^6; \qquad L = 63 \quad (14)$$

The resulting sequence $u_L(d)$ can be decomposed into elementary sequences $w_N(d)$, its shifted versions $T^jw_N(d)$ (with properly selected values of j) and null sequences, where $w_N(d)$ is generated by the polynomial

$$\phi_1(d) = 1 + d + d^3; \qquad N = 7 \quad (15)$$

The sequences $\{T^jw_N(d), j = 0, \ldots, N-1\}$ and their integer maps are shown in Table 1.

Table 1: Integer maps of two sets of elementary m-sequences generated by $\phi_1(d) = 1 + d + d^3$ and $\phi_2(d) = 1 + d^2 + d^3$

Elementary m-sequences of $\phi_1(d)$ Sequence	Integer map	Elementary m-sequences of $\phi_2(d)$ Sequence	Integer map
T^0w 1 1 1 0 1 0 0	0	T^0z 1 0 0 1 0 1 1	0
T^1w 1 1 0 1 0 0 1	1	T^1z 0 0 1 0 1 1 1	1
T^2w 1 0 1 0 0 1 1	2	T^2z 0 1 0 1 1 1 0	2
T^3w 0 1 0 0 1 1 1	3	T^3z 1 0 1 1 1 0 0	3
T^4w 1 0 0 1 1 1 0	4	T^4z 0 1 1 1 0 0 1	4
T^5w 0 0 1 1 1 0 1	5	T^5z 1 1 1 0 0 1 0	5
T^6w 0 1 1 1 0 1 0	6	T^6z 1 1 0 0 1 0 1	6

Here $S = 9$ and $P = 8$. Following the procedure of Reference 7, and using the integer representations of Table 1, the interleaving order of $w_N(d)$ for obtaining $u_L(d)$ is given by

$$I_P \triangleq \{-, 4, 0, 6, 6, 3, 4, 6, 4\} \tag{16}$$

where '−' represents the mapping of a position onto a 'null' sequence. A more complete expansion of u_L ($L = 63$) would read as follows:

$$\begin{aligned} u_L \triangleq \{&'-', 4, 0, 6, 6, 3, 4, 6, 4, '-', 5, 1, 0, 0, 4, 5, 0, 5,\\ &'-', 6, 2, 1, 1, 5, 6, 1, 6, '-', 0, 3, 2, 2, 6, 0, 2, 0,\\ &'-', 1, 4, 3, 3, 0, 1, 3, 1, '-', 2, 5, 4, 4, 1, 2, 4, 2,\\ &'-', 3, 6, 5, 5, 2, 3, 5, 3\} \end{aligned} \tag{17}$$

The complete expansion can be seen to have N (seven in this case) segments following each other. The S elements (nine in this case) of any segment, however, can be obtained simply by incrementing (modulo N) each of the corresponding elements of the previous segment. Any one of these segments is therefore sufficient to specify the integer series representation of the complete sequence $u_L(d)$. In Section 2.3 it will be seen that it is sufficient to expand out two successive segments of this representation for the calculation of the ACF of the composite sequence $u_L(d)$.

2.3 Autocorrelation function of composite m-sequences

Although the ACF of the longer (composite) m-sequence is well known, it is instructive to obtain it from its integer series representation in terms of component m-sequences, as was discussed in the preceding Sections. The motivation for this exercise here lies in the insight it yields in identifying the structure of other, not necessarily maximal length sequences, having an ACF identical to that of m-sequences.

A simple procedure to calculate the ACF of a composite (not necessarily maximal length) sequence based on the integer series representation is outlined in the form of a matrix computation in Table 2.† The $S \times 2S$ matrix (where S is the size of each of the N segments in $u_L(d)$) has its kth row and column designated by r_k and c_k, respectively, where

$$\begin{aligned} r_k = c_k &= \text{integer map of } k\text{th position of}\\ &\quad \text{sequence } u_L(d),\ k = 0, 1, \ldots, S-1\\ c_{k+S} &= c_k + 1 \end{aligned} \tag{18}$$

Note that several of the rows and columns may have the same designation, and $(S - P)$ of these will be designated by the symbol '−', corresponding to the mapping of a position onto a null sequence.

† The matrix in Table 2 can in fact have a dimension of $L \times L$ if one complete period of the sequence $u_L(d)$ is chosen for its construction. It is, however, sufficient to construct a smaller $S \times 2S$ matrix, based only on two segments of the sequence $u_L(d)$, to derive the necessary conclusions about the ACF properties of interest in this Section.

The entries e_{ij} of the matrix (representing the element in row i and column j) are now filled up as follows:

$$e_{ij} = (c_j - r_i) \bmod N,\ i, j = 0, 1, 2, \ldots \tag{19}$$

The following subtraction rules are used whenever r_i or c_j takes up a 'null' value '−':

$$(a - b) \bmod N \triangleq \begin{cases} \infty, \text{ when one of the two operands} \\ a \text{ and } b \text{ has a null value} \\ '-', \text{ when both } a \text{ and } b \text{ have} \\ \text{'null' values} \end{cases} \tag{20}$$

The value of e_{ij} represents the amount of shift required to make the subsequences r_i and c_j in the composite sequence $u_L(d)$ have coincident bit values (i.e. become identical in phase). The rules of eqn. 20 are therefore motivated by the fact that no amount of shifting of the column $u_L(d)$ with respect to its row counterpart would render r_i and c_j phase-coincident if one of these represents a null sequence and the other does not.

As illustrated in Table 2 for the example $u_L(d)$ of Section 2.2, the main diagonal thus has only zero and null value entries, indicating that all the subsequences have identical phases (and are therefore completely bit-coincident) when the row and column sequences have no mutual phase shift. In general, the jth diagonal will contain information regarding the relative phases of the subsequences in the corresponding positions of the j-shifted and unshifted composite sequences.

The matrix as constructed above contains important information regarding the ACF properties of the sequence $u_L(d)$ and also yields an understanding of the structure of alternate sequences having similar ACFs. These properties are taken up in the following remarks.

Remark 1: The jth diagonal, $j \neq 0$, has exactly $(S - P)$ zero or null entries for a given composite m-sequence $u_L(d)$.

Proof: A zero or null value in the ith position of the jth diagonal indicates that the subsequences in the ith positions of $u_L(d)$ and $T^j u_L(d)$ are identical. It therefore implies that the sum sequence

$$v^j(d) \triangleq u_L(d) + T^j u_L(d) \tag{21}$$

would also have a null subsequence at its ith position (in the integer map domain). Conversely, a null subsequence in any position of $v^j(d)$ would also imply a zero or null entry at the corresponding position on the jth diagonal.

However, $v^j(d)$ as defined above can be identified to be another shifted m-sequence having exactly $(S - P)$ null subsequences. It follows, therefore, that the jth diagonal has exactly $(S - P)$ null or zero entries.

Remark 2: The proof given above also shows that the various null values on the jth diagonal have a one-to-one correspondence with the null subsequences in $v^j(d)$, with each of the remaining entries on the diagonal (including ∞) corresponding to an m_1-subsequence. It must be

Table 2: Procedure for ACF calculation of composite 'm-derived sequences

u_L	—	4	0	6	6	3	4	6	4	—	5	1	0	0	4	5	0	.	.	.	.	.	.
—	—	∞	∞	∞	∞	∞	∞	∞	∞			∞											
4		0	3	2	2	6	0	2	0	∞			3										
0			0	6	6	3	4	6	4	∞	5			0									
6				0	0	4	5	0	5	∞	6	2			5								
6					0	4	5	0	5	∞	6	2	1			6							
3						0	1	3	1	∞	2	5	4	4			4						
4							0	2	0	∞	1	4	3	3	0			1					
6								0	5	∞	6	2	1	1	5	6			∞				
4									0	∞	1	4	3	3	0	1	3			2			
⋮										$I=j=0$									$I=j=8$		$I=11, j=2$		

emphasised here that an integer p in the diagonal of the matrix denotes the value of the relative phase between the concerned row and column subsequences, and not the subsequence $T^p w_N(d)$. This correspondence between the jth diagonal and $v^j(d)$, in turn, implies that both of them contain an identical distribution of ones and zeros, since both contain $(S - P)$ null entries and P m_1-subsequences.

Remark 3: The normalised ACF of $u_L(d)$ with lag j can be obtained from the well known relation

$$\theta(j) = 1 - 2P_j(1) \tag{22}$$

where $P_j(1)$ is the probability of a one in $v^j(d)$, or, as was discussed above, in the binary version of the jth diagonal sequence. Since the latter contains $(S - P)$ null entries and P m_1-subsequences, it can be said that the number of ones in $v^j(d)$ is given by

$$L_1 = L - (S - P)N - \frac{P}{2}(N - 1) = \frac{P(N + 1)}{2} \tag{23}$$

as each m_1-subsequence contains $(N - 1)/2$ zeros.

It may be noted that eqn. 23 also follows trivially from the fact that $v^j(d)$ itself is a composite m-sequence having a similar structure. However, our emphasis here is on obtaining L_1 from the diagonal sequence rather than $v^j(d)$ itself for exploitation in the construction of new sequences, as in Section 3.

It follows from eqn. 23 that

$$P_j(1) = \frac{2(N + 1)}{2L} \tag{24}$$

so that

$$\theta(j) = 1 - \frac{L + 1}{L} = -\frac{1}{L} \tag{25}$$

as expected for an m-sequence of length L.

Remark 4: Although the sequence u contains N segments, each of length S, it suffices to obtain the first S diagonals, since all the rest can be obtained from these in a simple manner. Thus, for the lth diagonal, $l = nS + j$, we have

$$e_{il} = e_{ij} + n(\text{mod } N),\ l \in (0, 1, \ldots,),\ j \in (0, 1, \ldots, S - 1) \tag{26}$$

with the rules $\infty + n \triangleq \infty$ and '$-$' $+ n \triangleq$ '$-$', as before. This is illustrated for $l = 11$ in Table 2.

3 Construction of *m*-like interleaved cipher sequences

An important conclusion that emerges from the ACF studies of Section 2 is that the ACF $\theta_u(\cdot)$ of a composite sequence $u(d)$, consisting of S interleaved m_1-subsequences each of length N, depends only on the structure of the diagonals of the matrix comprising the modulo-N difference, or relative phases of the subsequences in the shifted and unshifted composite sequences $u(d)$ and $T^j u(d)$. Note, however, that the properties of the diagonal sequences considered in Section 2 are invariant to the actual choice of the m_1-subsequences $w_N(d)$, as long as the chosen subsequence, say, $z_N(d)$, is interleaved in the same manner as in the composite m-sequence, i.e. using the same interleaving order I_P. This invariance is a result of the fact that such a modification leaves the integer series representation of the new sequence, say, $u'(d)$, unaltered.

It therefore follows that, if we replace the subsequences $T^j w_N(d)$ (chosen earlier so as to render $u_L(d)$ an m-sequence of length L) by alternative m_1-subsequences $T^j z_N(d)$ and interleave these along with null sequences as for $u_L(d)$, the resulting composite sequence $u'(d)$ will yield an identical diagonal structure to that of $u_L(d)$, and will therefore have ACF properties identical to that of $u_L(d)$, even though $u'(d)$ may no longer be an m-sequence. It is these sequences $u'(d)$ which we call here 'm-like cipher sequences'. The following theorem highlights the important properties of the composite sequences u'.

Theorem: The composite sequence u', having the same integer series representation as the m-sequence $u_L(d)$ but obtained by replacing the elementary subsequences $T^j w$ with $T^j z$, where z is a different m_1-subsequence of length N, has the following properties:

$$(a)\ \theta_{u'}(j) = -1/L \qquad j \neq 0 \tag{27}$$

$$(b)\ L_{u'} = \text{lengthy or periodicity of } u'(d) = S \cdot N \tag{28}$$

$$(c)\ L_1(u') = \text{number of ones in } u'(d) = (L + 1)/2 \tag{29}$$

$$(d)\ u'(d) \text{ is not an } m\text{-sequence.} \tag{30}$$

Proof:

(*a*) This, as discussed above, follows from the fact that the diagonals of Table 2 remain unchanged if $T^j w$ is replaced with $T^j z$, where z is an m_1-subsequence of the same length as w. The ACF therefore has the same properties as any m-sequence of length L

(*b*) $L_{u'} = S \cdot N$ by construction

(*c*) $L_1(u') = P((N + 1)/2) = (L + 1)/2$ by construction; therefore $L_0(u') = (L - 1)/2$

(*d*) The premise that the sequence $u'(d)$ is not an m-sequence is based on the fact that there exists a unique interleaving order I_P for which the composite sequence in terms of $T^j w$ (eqns. 13 and 6) is an m-sequence [7]. When w is replaced by z, the corresponding interleaving order I'_P required to produce the longer m-sequence would therefore be different, i.e. $I'_P \neq I_P$ [7]. Since, in our construction of $u'(d)$, the old interleaving order I_P is employed, it follows that $u'(d)$ is not an m-sequence. Two examples of m-like sequences and their construction procedure are presented in Table 3 for illustration.

Table 3: Construction of *m*-like sequences

Composite sequences	Subsequences	Interleaving order
1 m-sequence u	w	{−4 0 6 6 3 4 6 4} or {**0**, T^4w, T^0w, T^6w, T^6w, T^3w, T^4w, T^6w, T^4w}
2 m-sequence v	z	{6 4 6 0 4 4 3 6 −} or {T^6z, T^4z, T^6z, T^0z, T^4z, T^4z, T^3z, T^6z, **0**}
3 m-like sequence u'	z	{**0**, T^4z, T^0z, T^6z, T^6z, T^3z, T^4z, T^6z, T^4w}
4 m-like sequence v'	w	{T^6w, T^4w, T^6w, T^0w, T^4w, T^4w, T^3w, T^6w, **0**}

5 m-like sequences in binary form

u': **0**0111101**0**·**0**1011011**1**·**0**1000110**1**·
01100110**1**·**0**0011101**0**·**0**0100000**0**·
01111011**1**

v': 0101110**00**·1011001**10**·1011000**10**·
1110110**10**·0101111**00**·1110111**10**·
0000010**00**.

4 Predictability of *m*-like interleaved cipher sequences

It is well known that it is possible to identify an entire m-sequence after deciphering $2m$ consecutive bits of the sequence. This arises because of the linear recursive relation that exists among the bit patterns comprising an m-sequence. This can also be appreciated from the fact that it

is only necessary to identify the mth-order generating polynomial and the m-bit initial state of the shift register to completely define an m-sequence.

The story, however, is quite different for the interleaved sequence $u'(d)$ constructed in Section 3. Observation of $2m$ consecutive bits will not help here, since $u'(d)$ is not an m-sequence. The identification of the sequence $u'(d)$ will now require the implementation of the following procedure.

Assume first that the values of L, S and N are somehow known to an interceptor. The degree m_1 of the characteristic polynomial generating the m_1-sequence $z(d)$ is therefore also known. In order to identify the subsequence $z(d)$, it is once again necessary to decode its $2m_1$ consecutive bit values from the message. These bits are, however, dispersed over $2m_1S$ consecutive bits of the sequence $u'(d)$. It is, therefore, necessary now to decipher $2m_1S$ consecutive bits and then decimate them by a factor S in order to identify $z(d)$ and hence $u'(d)$. This procedure assumes that the interleaving order I_P is also known to the interceptor. Even so, the complexity of the sequence $u'(d)$ can be said to be of $2m_1S$ bits against $2m$ bits of the m-sequence $u_L(d)$. For long sequences

$$\frac{2m_1S}{2m} \gg 1$$

so that the interceptor now has to correctly decipher and store a much larger number of bits before he can hope to identify the sequence $u'(d)$.

This estimate of complexity is, in fact, somewhat pessimistic, in that it assumes prior knowledge of the values of L, S and N. In practice, such knowledge is rarely justifiable. In fact, for a given large value of L, there are many ways of choosing S and L with $L = S \cdot N$, along with corresponding different choices of the subsequences z of length N. Similarly, the knowledge of interleaving order I_P assumed in the above procedure adds to the pessimistic nature of the above estimate of complexity.

To appreciate the magnitude of increased complexity, consider the case with $m = 12$, so that

$$L = 2^{12} - 1 = 4095 \text{ bits}$$

Two choices for S and N are as follows:

(*a*) $S = 65$, $N = 63$; in this case the ratio $m_1S/m = 32.5$ and the value for $2m_1S = 780$ bits

(*b*) $S = 237$ and $N = 15$; in this case the ratio $m_1S/m_1 = 91$ and the value of $2m_1S = 2184$ bits.

Thus, in (*a*) the interceptor has to obtain and store 780 bits of $u'(d)$ from the ciphered text and in (*b*) 2184 bits from the ciphered text, as compared to the 24 bits required for identifying $u_L(d)$. Note that the value of $2m_1S \simeq (1/2)L$ in case (*b*).

This argument and the above example only serve to illustrate the increased complexity associated with the m-like sequences. An appropriate measure of the complexity of nonlinear sequences is the so-called 'equivalent span' (ELS), which may be defined to be the least length of a linear feedback shift register which can produce the sequence of interest [8]. This quantity, however, is usually difficult to compute, and the value m_1S used above is only an upper bound of the ELS value for the m-like sequences. The exact estimation of the complexity of these sequences via the calculation of the ELS is still being considered [9].

5 Conclusion

A new class of cipher sequences, called 'm-like cipher sequences', are introduced which have the same ACF properties as an m-sequence, but which are much more complex to decode from partial observation. The construction of the m-like cipher sequences is based on interleaving elementary m-sequences and null sequences in a particular order.

6 References

1 KEY, E.L.: 'An analysis of the structure and complexity of nonlinear binary sequence generators', *IEEE Trans.*, 1976, **IT-22**, pp. 732–736

2 JENNINGS, S.M.: 'Autocorrelation function of the multiplexed sequence', *IEE Proc. F, Commun., Radar & Signal Process.*, 1984, **131**, (2), pp. 169–172

3 OLSEN, J.D., SHOLTZ, R.A., and WELCH, L.R.: 'Bent function sequences', *IEEE Trans.*, 1982, **IT-28**, pp. 856–864

4 WEINRICHTER, H., and SURBÖCK, F.: 'Phase normalised m-sequences with the inphase decimation property $\{m(k)\} = \{m(2k)\}$', *Electron. Lett.*, 1976, **12**, pp. 590–591

5 QUYNH, L.C., and PRASAD, S.: 'On the autonomous response of cascaded scramblers and interleaved m-sequences', *IEEE Trans.* (under consideration)

6 MCWILLIAMS, F.J., and SLOANE, N.J.: 'Psuedo-random sequences and arrays', *Proc. IEEE*, 1976, **64**, pp. 1715–1729

7 SURBOCK, F., and WEINRICHTER, H.: 'Interlacing properties of shift register sequences with generator polynomials irreducible over $GF(p)$', *IEEE Trans.*, 1978, **IT-24**, pp. 386–389

8 MASSEY, J.L.: 'Shift register synthesis and BCH decoding', *ibid.*, 1969, **IT-15**, pp. 122–129

9 QUYNH, L.C., and PRASAD, S.: 'On the equivalent linear span analysis of nonlinear binary sequences having an interleaved structure', *IEE Proc. F* (under consideration)