ON QUADRATIC M-SEQUENCES*

A.H. Chan[1], R.A. Games[2] and J.J. Rushanan[2]

[1] College of Computer Science, Northeastern University, Boston, MA 02115
[2] The MITRE Corporation, Bedford, MA 01730

1 Introduction

Maximal sequences generated by linear feedback shift registers (FSRs), known as m-sequences, have been well-studied in the literature [1]. These sequences have long period, good statistical properties and two-valued autocorrelation functions. They have been used as pseudorandom sequences in simulations, in spread spectrum communications, and as test sequences. However, due to the linearity properties exhibited in these sequences, m-sequences are extremely vulnerable to Known Plaintext Attack by cryptanalysts. Specifically, a binary m-sequence of span n can be completely determined by the Berlekamp-Massey algorithm after $2n$ bits of the sequence are observed.

In order to overcome the weakness found in m-sequences, nonlinearities have been introduced. Such nonlinearities include the use of a nonlinear feedforward function on a linear FSR [2, 3] and the use of a nonlinear combining function on several linear FSRs [4, 5]. However, not much has been studied with respect to the use of nonlinear feedback functions. This is partly due to the fact that the study of nonlinear functions proves to be a formidable task.

In this paper, we study nonlinear feedback functions by first investigating quadratic functions. The *quadratic span* of a periodic binary sequence is the length of the shortest quadratic FSR that generates the sequence. Previously, bounds on the quadratic spans of full cycle DeBruijn sequences of span n were obtained in [6], which also partially generalized the Berlekamp-Massey algorithm to the quadratic case. The lower bound on the quadratic span of a DeBruijn sequence of span n was improved to $n + 2$ in [7]. By eliminating the all zero state from the DeBruijn sequence, the period is reduced from 2^n to $2^n - 1$. This paper considers the question as to whether the resulting sequence can now have quadratic span n. Such sequences are the quadratic analog of the linear m-sequences and present an attractive extremal case to explore further the structure of nonlinear FSRs.

Definition 1. A binary sequence $\mathbf{s} = s_0, s_1, \ldots$ satisfies a *quadratic recurrence*

* A.H. Chan is supported by the National Security Agency under Grant Number MDA904-91-H-0012 and partially supported by the MITRE Sponsored Research Program. R.A. Games and J.J. Rushanan are supported by the MITRE Sponsored Research Program. The United States Government is authorized to reproduce and distribute reprints notwithstanding any copyright notation hereon.

of span n if for all $k \geq 0$,

$$s_{n+k} = \sum_{0 \leq i \leq j \leq n-1} a_{ij} s_{i+k} s_{j+k}, \tag{1}$$

where $a_{ij} = 0$ or 1.

Definition 2. A binary sequence s is said to be a *quadratic m-sequence of span* n if it satisfies a quadratic recurrence of span n and has period $2^n - 1$.

For every sequence s with recurrence relation as shown in (1), there associates a quadratic feedback function f on n variables, given by

$$f(x_0, \ldots, x_{n-1}) = \sum_{0 \leq i \leq j \leq n-1} a_{ij} x_i x_j \tag{2}$$

We note that $x_i^2 = x_i$ for all $i \geq 0$ over GF(2), thus the coefficients a_{ii} correspond to the linear terms in f. In section 2, we study quadratic m-sequences by considering the algebraic structures of the feedback functions given in (2). In section 3, we consider algorithmic generation of quadratic m-sequences from linear m-sequences.

2 Quadratic Feedback Functions

For any initial loading (state) of a feedback shift register, the state updates as each sequence bit is produced. Since there are only finitely many states, the states have to repeat eventually. If the initial state is never repeated, the cycle of states generated is said to have a branch point. For quadratic m-sequences, there are exactly two cycles without branch points, namely the all zero state cycle and the cycle consisting of all other $2^n - 1$ nonzero states. In [1], it is proved for general feedback functions that

Theorem 3. *The cycles generated by a feedback shift register have no branch points if and only if its feedback function can be decomposed as*

$$f(x_0, \ldots, x_{n-1}) = x_0 + g(x_1, \ldots, x_{n-1}).$$

Corollary 4. *Let f be a feedback function that generates a quadratic m-sequence. Then*

$$f(x_0, \ldots, x_{n-1}) = x_0 + g(x_1, \ldots, x_{n-1}), \tag{3}$$

where g is a quadratic function.

Thus, to study $f(x_0, \ldots, x_{n-1})$ we consider $g(x_1, \ldots, x_{n-1})$ instead. Let $n_L(g)$ and $n_Q(g)$ denote respectively the number of linear and quadratic terms in g. We show that

Theorem 5. *If $x_0 + g(x_1, \ldots, x_{n-1})$ generates a quadratic m-sequence, then*

$$n_L(g) + n_Q(g) \equiv 1 \mod 2.$$

Proof. Consider the all-ones state corresponding to each x_i having the value 1. If the next element of the sequence determined by the feedback function in equation (3) were a 1, then the state of the register would not change and we we would have a cycle consisting of n 1s. Since this clearly cannot happen, we must have the next state equal a zero, that is,

$$1 + g(1, \ldots, 1) \equiv 0 \qquad \bmod 2.$$

The value of $g(1, \ldots, 1)$ is the same modulo 2 as the the sum of the number of linear terms and the number of quadratic terms, $n_L(g) + n_Q(g)$, which completes the result. ▮

If we consider the special term x_0 as a linear term, then theorem 5 says that for a feedback function generating a quadratic m-sequence, the number of linear terms and the number of quadratic terms have equal parity. The next result shows that there must be a linear term besides x_0.

Theorem 6. *If $x_0 + g(x_1, \ldots, x_{n-1})$ generates a quadratic m-sequence, then $n_L(g) \neq 0$, equivalently, $g(x_1, \ldots, x_{n-1})$ must contain some linear term.*

Proof. Suppose that g has no linear terms and consider the state with $x_0 = 1$ and the other $x_i = 0$. The next state, regardless of g, has $x_{n-1} = 1$ and the rest of the $x_i = 0$. Since g has no linear terms, the next state has all zeros except for $x_{n-2} = 1$. Continuing, we see that after n steps, we are back to the initial state, which is a contradiction. ▮

If a sequence s has maximum period, then its reverse, **R**(s), also has maximum period. Thus, we have the following,

Theorem 7. *If $x_0 + g(x_1, \ldots, x_{n-1})$ generates a quadratic m-sequence, then $x_0 + g(x_{n-1}, \ldots, x_1)$ generates a quadratic m-sequence.*

Proof. If

$$x_n = x_0 + g(x_1, \ldots, x_{n-1})$$

then

$$x_0 = x_n + g(x_1, \ldots, x_{n-1}).$$

If we let $y_i = x_{n-i}$, then the latter equation becomes

$$y_n = y_0 + g(y_{n-1}, \ldots, y_1),$$

which is the recurrence relation for the reversal sequence. ▮

We remark that although we stated theorems 5, 6, and 7 for quadratic feedback functions, their generalizations to arbitrary nonlinear functions are straightforward.

3 Generation of Quadratic m-Sequences

In this section, we consider the generation of quadratic m-sequences by introducing quadratic terms to the feedback function of an m-sequence. Because the number of linear terms in a primitive polynomial is odd (and so the number of linear terms in the feedback function is even), theorem 5 implies that the introduction of a quadratic term requires the "addition" of a linear term. However, the extra linear term that is added may cancel with an existing linear term in the feedback function; thus it effectively could reduce the number of linear terms in the feedback function.

To study the generation of quadratic m-sequence, we start by considering the simplest device that can be added to the feedback function of an m-sequence. Such a device corresponds to the addition of terms $x_i + x_i x_j$ to the feedback function for $i \neq j$ and $i \neq 0 \neq j$. This device affects the state changes in the FSR if and only if $x_i = 1$ and $x_j = 0$. Let $v_1 = (v_{1,0}, \ldots, v_{1,n-1})$ be the first state where $v_{1,i} = 1$ and $v_{1,j} = 0$. Due to the addition of the device, the state change differs from that of the m-sequence at v_1 and v_1', where $v_1' = (1 + v_{1,0}, v_{1,1}, \ldots, v_{1,n-1})$. This results in the cycle of states being broken into two separate cycles, as shown in figure 1. Here the cycle on the left represents the m-sequence with dots showing successive states; the resulting two cycles are shown on the right. As states $v_\ell = (v_{\ell,0}, \ldots, v_{\ell,n-1})$, where $v_{\ell,i} = 0, v_{\ell,j} = 1$, are encountered, the cycles may be further broken up or joined together, depending on whether v_ℓ and v_ℓ' are on the same cycle or not.

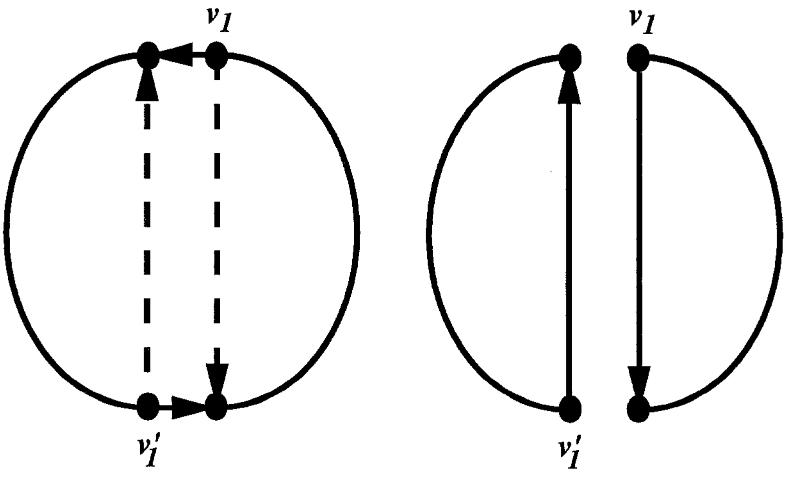

Figure 1. Decomposition of an m-Sequence into Two Cycles

For fixed $i \neq j$, there are exactly 2^{n-3} state pairs (v_ℓ, v'_ℓ), where $v_{\ell,i} = 1$ and $v_{\ell,j} = 0$. The sequence of appearances of $v_1, v_2, \ldots, v_{2^{n-3}}, v'_1, v'_2, \ldots, v'_{2^{n-3}}$ in the m-sequence cycle is defined to be the *pattern of state changes* for the device $x_i + x_i x_j$. The appearances of v_i's have to be in ascending order; however, the v'_i need not be in ascending order.

Let m be the length of a pattern of state changes, so necessarily m is even.

Definition 8. A pattern of state changes $P = (p_1, \ldots, p_m)$ is said to be a *legal pattern* if

1. if $p_i = v_\ell, p_j = v'_\ell$, then $i < j$,
2. if $p_i = v_\ell, p_j = v_k$ and $l < k$, then $i < j$.

Proposition 9. *The number of legal patterns of length m is exactly* $1{\cdot}3{\cdot}5 \cdots (m-1)$.

Proof. Let $r = m/2$. We count how many ways there are to place v'_i for $i = 1, \ldots, r$ in the ascending sequence $v_1, \ldots, v_r$. There is only one place to put v'_r, namely to the right of v_r. There are then three places to place v'_{r-1}, namely between v_{r-1} and v_r, between v_r and v'_r, and to the right of v'_r. Once v'_{r-1} is placed, there are now five possible positions for v'_{r-2}. Continuing in this manner yields $m-1$ positions to place v'_1. This completes the result. ∎

Obviously not every legal pattern of state changes produces a single cycle. A legal pattern of state changes is said to be *maximum* if it results in a single cycle. Necessarily, if the length of a maximum pattern is $m = 2r$, then r itself must be even; this is because every v_i, v'_i pair either splits a cycle into two or joins two cycles. Our goal is to characterize maximum legal patterns; below are some partial results. If adding a device to an m-sequence yields a maximum legal pattern, then we have obtained a quadratic m-sequence. In such a case we have $m = 2^{n-2}$.

Proposition 10. *If P is a maximum pattern of length $m = 2r$, then for all $i, 1 \leq i \leq r, p_i = v_\ell$ implies $p_{i+1} \neq v'_\ell$.*

Proof. If v_ℓ and v'_ℓ are consecutive in the pattern, then they partition the cycle into two pieces, one of which can never be joined to the other. ∎

Proposition 11. *[8] Suppose that P is a maximal legal pattern. If for some i, v_i and v_{i+1} occur consecutively in P, then v'_{i+1} and v'_i cannot occur consecutively.*

Proof. Suppose that both pairs (v_i, v_{i+1}) and (v'_{i+1}, v'_i) occur in P. Then there is the following cycle, which violates P being maximum;

$$\mathrm{succ}(v_i) \rightarrow v_{i+1} \rightarrow \mathrm{succ}(v'_{i+1}) \rightarrow v'_i \rightarrow \mathrm{succ}(v_i),$$

where $\mathrm{succ}(v_i)$ denotes the successor to v_i in the original cycle and an arrow denotes a path through all intervening states (which cannot contain any of the v_k). ∎

Proposition 12. *The pattern*

$$v_1\ v_2\ \ldots\ v_r\ v_1'\ v_2'\ \ldots\ v_r'$$

is maximum if r is even.

Proof. We illustrate the complete cycle, using the notation given in the proof of proposition 11. The cycle begins as

$$\begin{array}{ll} & v_1 \to \mathrm{succ}(v_1') \to v_2' \to \mathrm{succ}(v_2) \\ \to & v_3 \to \mathrm{succ}(v_3') \to v_4' \to \mathrm{succ}(v_4) \\ & \vdots \\ \to & v_{r-1} \to \mathrm{succ}(v_{r-1}') \to v_r' \to \mathrm{succ}(v_r) \end{array} \tag{4}$$

Notice that we need r to be even. Now the next state change after v_r is v_1', so the cycle continues as

$$\begin{array}{ll} \to & v_1' \to \mathrm{succ}(v_1) \to v_2 \to \mathrm{succ}(v_2') \\ \to & v_3' \to \mathrm{succ}(v_3) \to v_4 \to \mathrm{succ}(v_4') \\ & \vdots \\ \to & v_{r-1}' \to \mathrm{succ}(v_{r-1}) \to v_r \to \mathrm{succ}(v_r') \\ \to & v_1 \end{array} \tag{5}$$

We used that the next state change after v_r' is v_1. To finish the proof, it is easy to see that every possible sequence of states occurs in the above cycle. ▮

We hope to establish the existence of quadratic m-sequences for all spans by showing that the addition of a device of the form $x_i + x_i x_j$ to some primitive feedback polynomial will generate a quadratic m-sequence for all $n \geq 4$. We discuss some experimental results of this technique in the next section.

4 Enumeration of Quadratic m-Sequences

Our first table enumerates all quadratic m-sequences of span n, $n \leq 7$. These numbers include the linear m-sequences; in particular, there are no non-linear quadratic m-sequence for $n < 4$.

n	1	2	3	4	5	6	7
Quadratic m-Seq.	1	1	2	16	128	1952	64,056

We enumerated quadratic m-sequences of a special form, namely those with as few taps as possible: one quadratic term and two linear terms (besides x_0). The table below shows the number of quadratic m-sequences of the indicated form (the x_0 is not included). The first form includes the second form. We extended the enumeration of the second form up to $n = 18$, but there were not any more such quadratic m-sequences. The fact that the number of quadratic m-sequences of these forms decreases (to zero, in the case of the second form) leaves us to believe that these forms are not a promising avenue to pursue in order to generate infinite families of quadratic m-sequences.

n	4	5	6	7	8	9	10	11	12
$x_i + x_j + x_k x_\ell$	6	8	16	14	30	8	14	6	6
$x_i + x_j + x_i x_j$	2	2	6	2	4	0	0	0	2

Finally, we show the results of adding a device of the form $x_i + x_i x_j$ to an m-sequence (where $i \neq j$ and $i \neq 0 \neq j$). In the following table, we give the number of linear m-sequences and the number of linear m-sequences that can be extended by some device. The last row shows the total number of quadratic m-sequences that result from a device added to an linear m-sequence; note that the same m-sequence can sometimes be extended by several different devices. Thus, adding a device to a linear m-sequence is a promising approach to generate quadratic m-sequences.

n	4	5	6	7	8	9	10	11
Linear m-Seq.	2	6	6	18	16	48	60	176
Extendable	2	6	6	14	16	20	28	50
Quadratic m-Seq.	4	10	12	24	20	24	34	56

5 Future Work

This paper presents a first step in the investigation of nonlinear feedback shift register sequences. We have concentrated on the construction of quadratic m-sequences by adding a simple device of the form $x_i + x_i x_j$ to a linear m-sequence (where $i \neq j$ and $i \neq 0 \neq j$). We conjecture from our study that

Conjecture 1 *For each $n \geq 4$, there exists a linear feedback function f that generates an m-sequence and integers i, j with $i \neq j$ and $i \neq 0 \neq j$ such that*

$$f(x_0, x_1, \ldots, x_{n-1}) + x_i + x_i x_j \qquad (6)$$

generates a quadratic m-sequence.

One of the major problems to be considered is the characterization of maximum legal patterns. We have established some necessary conditions and are continuing in our search for sufficient conditions. We plan to extend the procedure of adding a device to a linear m-sequence to adding it to a quadratic m-sequence. In this way, we hope to recursively generate all quadratic m-sequences.

References

1. S.W. Golomb, *Shift Register Sequences*, Aegean Park Press, California, 1982.
2. B. Gordon, W.H. Mills and L.R. Welch, *Some New Difference Sets*, Canad. J. Math. 14 (1962), pages 614-625.
3. A.H. Chan and R.A. Games, *On the Linear Span of Binary Sequences Obtained from Finite Geometry*, IEEE Trans. on Information Theory (1990), vol IT-36, pages 548-552.
4. E.L. Key, *An Analysis of the Structure and Complexity of Nonlinear Binary Sequence Generators*, IEEE Trans. on Information Theory (1976), Vol IT-22.
5. R.A. Rueppel and O.J. Staffelbach, *Products of Linear Recurring Sequences with Maximum Complexity*, IEEE Trans. on Information Theory (1987), vol IT-33, pages 124-129.
6. A.H. Chan and R.A. Games, *On the Quadratic Spans of DeBruijn Sequences*, IEEE Trans. on Information Theory (1990), vol IT-36 pages 822–829.
7. L.H. Khachatrian, *The Lower Bound of the Quadratic Spans of DeBruijn Sequences*, Designs, Codes, and Cryptography (1993), vol 3.
8. B. Sroka, personal communication