Lengthening m-sequences with shift sequences

Kirsten Nelson, Carleton University

September 29, 2017

Under the supervision of: Daniel Panario, Carleton University
Brett Stevens, Carleton University

1. Introduction
2. Interleaved Sequence with Periods $\leq s \cdot t$
3. Shift Sequence equivalence
4. Lengthening m-sequences
5. Coverage Properties of Interleaved Sequences

m-sequences over finite fields

Definition

An *m-sequence* over a finite field $\mathbb{F}_q$ is a sequence that satisfies Golomb's three randomness properties, as given in Golomb & Gong (2005):

- balance; the number of symbols varies by no more than one
- run; $1/q$ of the runs have length 1, $1/q^2$ have length 2, etc.
- and, the auto-correlation is two-valued.

An example of a sequence over $\mathbb{F}_2$ is the sequence $0, 0, 1, 0, 1, 1, 1, \ldots$. We can see that it is balanced, since there are four symbols of '1' and three symbols of '0'. The runs are 00, 1, 0, and 111, so it is true that $1/2$ of the runs have length 1 and $1/4$ have length 2.

Shift Sequences

The notion of a shift sequence was introduced by R. A. Games (1985). We use the definition given in He, Panario, & Wang (2010).

Definition

Given an l-ary sequence $\underline{a} = (a_0, \ldots, a_{s-1})$ of period s called the *base* sequence, we define the *shift sequence* to be a sequence $\underline{e} = (e_0, \ldots, e_{t-1})$, for each $0 \leq i \leq t-1$ such that $e_i \in \mathbb{Z}_s \cup \infty$.

This is not to be confused with the left shift operator, which we will also use:

Definition

For $\underline{a} = (a_0, a_1, \cdots)$, we define a left shift operator L as follows: $L^i \underline{a} = (a_i, a_{i+1}, a_{i+2}, \cdots)$.

Interleaved Sequences

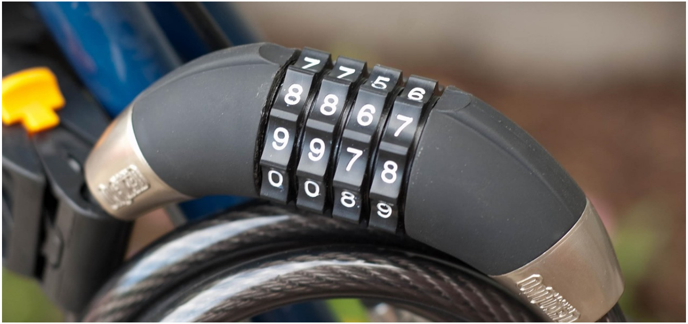
Definition

We combine the base sequence and shift sequence to create an *interleaved* sequence as follows: the sequence $\underline{u} = (u_0, \dots, u_{st-1})$ is an l-ary sequence of period $s \cdot t$. We arrange the elements of the sequence $\underline{u}$ into an $s \times t$ matrix as follows:

$$\mathcal{A}_u = \begin{bmatrix} u_0 & \cdots & u_{t-1} \\ \vdots & \vdots & \vdots \\ u_{(s-1)t} & \cdots & u_{(s-1)t+t-1} \end{bmatrix}$$

Each column of $\mathcal{A}_u$ is a shift of the base sequence $\underline{a}$. Let A_j be the j^{th} column. Then $A_j = L^{e_j}(\underline{a})$ and $L^{\infty}(\underline{a}) = (0, \dots, 0)$. The matrix $\mathcal{A}_u$ is called the *matrix form of sequence* $\underline{u}$, and $\underline{u}$ is called an interleaved sequence from the base sequence $\underline{a}$ and the shift sequence $\underline{e}$.

Example 0



The base sequence $\underline{a} = (0, 1, \ldots 9)$ is a sequence over $\mathbb{F}_{10}$ of length $s = 10$. The shift sequence $\underline{e} = [7, 7, 5, 6]$ shown is a sequence of length $t = 4$ with the elements taken from $\mathbb{Z}_{10}$. The interleaved sequence is $\underline{u} = (7, 7, 5, 6, 8, 8, 6, 7, 9, 9, 7, 8, 0, 0, 8, 9, 1, 1, 9, 0, \ldots)$ of length 40.

Picture courtesy of: wonderhowto.com

Example 1

In the definition we allow a shift sequence entry of ∞, which represents a column of the matrix consisting of all zeroes.

Let our base sequence be $\underline{a} = (0, 0, 1, 0, 1, 1, 1, \dots)$ and our shift sequence be $\underline{e} = [0, 1, 3, \infty]$. The matrix form of the interleaved sequence is:

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

The interleaved sequence is $\underline{u} = 0000011010100110110011001010\dots$.

Period of $\underline{u}$

If we begin with base sequence $\underline{a}$ with period s and shift sequence $\underline{e}$ with period t, when is the period of $\underline{u}$ equal to $s \cdot t$, and when is it shorter?

We know from Gong (1995) that the period of the interleaved sequence always divides st. If s and t are co-prime, then, the period of the interleaved sequence is either 1 (which can only happen if $\underline{a}$ has period 1), divides s, divides t, or has period st.

Example 2

Let the base sequence be $\underline{a} = (0, 1, 2, 3, 0, 1, 2, 3 \ldots)$ over $\mathbb{F}_4$. Consider the shift sequence $\underline{e} = [0, 3, 2]$. This creates the interleaved sequence $\underline{u} = 032103210321\ldots$, which has period 4.

$$\mathcal{A} = \begin{bmatrix} 0 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$$

Note that $s = 4$, $t = 3$, $s^t = 64$, and $s \cdot t = 12$.

Operations on shift sequences

We've identified two operations on shift sequences that produce shift-equivalent sequences.

- Increment each element of the shift sequence by one; this shifts the interleaved sequence by one row (number of characters shifted is t, the size of the row)
- Increment the first element of the shift sequence and move it to the end; this shifts the interleaved sequence by one character
- Side note: If the sequence is not palindromic, there will be a shift that creates a sequence that's a reverse of another shift

Example of shifting by a row

Here is the previous example. Base sequence $\underline{a} = (0, 0, 1, 0, 1, 1, 1 \dots)$. We compare the shift sequence $[0, 1, 3, \infty]$ to a shift sequence of $[1, 2, 4, \infty]$. We use the convention that $\infty + n = \infty$, and do our other additions modulo s, since entries in the shift sequence are in $\mathbb{Z}_s$.

$$\mathcal{A}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \mathcal{A}_2 = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

If $\underline{e}_1 = [e_0, e_1, \dots e_{t-1}]$ and $\underline{e}_2 = [e_0 + d, e_1 + d, \dots e_{t-1} + d]$, then we can generalize to a shift of d rows: $\underline{u}_2 = L^{td} \underline{u}_1$.

Example of shifting by a single character

The base sequence is again $\underline{a} = (0, 0, 1, 0, 1, 1, 1, \ldots)$. We compare the shift sequence $[0, 1, 3, \infty]$ to a shift sequence of $[1, 3, \infty, 1]$.

$$\mathcal{A}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \mathcal{A}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Assuming that $r < t$, $\underline{e}_1 = (e_0, e_1, \ldots e_{t-1})$ as before, and $\underline{e}_3 = (e_r, e_{r+1}, \ldots e_{t-1}, e_0 + 1, e_1 + 1, \cdots e_{r-1} + 1)$, then we can express this as: $\underline{u}_3 = L^r(\underline{u}_1)$

Overall equivalence

We combine these so that any shift $n \in \mathbb{Z}$ can be represented. Let $n = td + r$ where $r < t$. Our desired shift is $L^n = L^{td+r}$, and is implemented by modifying the shift sequence this way:

$\underline{e}_n = (e_r + d, e_{r+1} + d, \cdots, e_{t-1} + d, e_0 + d + 1, e_1 + d + 1, \cdots, e_{r-1} + d + 1)$

The canonical shifts

This leads to two criteria to define our canonical shift sequences.

- The shift sequence must begin with the zero symbol
- The symbols must be the lexicographically smallest formation through rotations

Example

Consider our previous example where the base sequence was $\underline{a} = (0, 1, 2, 3 \dots)$ over $\mathbb{F}_4$. We don't actually need any of the values of the base sequence; we only need to know the length s. These are our canonical shifts.

- $[0, 0, 0]$
- $[0, 0, 2]$
- $[0, 0, 3]$
- $[0, 1, 0]$
- $[0, 1, 3]$
- $[0, 3, 2]$

Recall that there are $s^t = 4^3 = 64$ possible shift sequences.

Example

Let's consider the first shift, $[0, 0, 0]$. We can cycle through all the possible shift-equivalent interleaved sequences associated with it.
$[0, 0, 0] \rightarrow [0, 0, 1] \rightarrow [0, 1, 1] \rightarrow [1, 1, 1] \rightarrow [1, 1, 2] \rightarrow [1, 2, 2] \cdots$
The pattern is obvious from here, and since it does not repeat after six single-character rotations, it must continue to twelve.
The first five sequences work the same way, representing $5 \times 12 = 60$ of the possible 64 shift sequences.

This shift creates a sequence of length 4:
$[0, 3, 2] \rightarrow [3, 2, 1] \rightarrow [2, 1, 0] \rightarrow [1, 0, 3]$
This accounts for the final four shift sequences.

Shift sequence entries of ∞

We also have a provision to have entries in our shift sequence of ∞, which translate to a column of zeroes. We'll restrict ourselves to one such entry in this talk, and since shift sequences can be rotated as we've seen above, we'll assume it's the first entry when it exists.

These are important to 'repair' the balance when we go from a shorter *m*-sequence to a longer one.

Decomposing m-sequences

It is well-known (e.g. Golomb and Gong, 2005) that any *m*-sequence can be written as an interleaved sequence, unless the period is a prime number. Example: Starting with the binary *m*-sequence $\underline{u} = 000100110101111$ with period 15, we can write it as an interleaved sequence by arranging it in a 3 by 5 array:

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

From this we see that the appropriate shift sequence is $\underline{u} = [\infty, 0, 0, 2, 0]$.

Degree of sequences

The choice of 3 by 5 is forced by the relative lengths of the *m*-sequences used.
Recall that if the degree of an *m*-sequence over $\mathbb{F}_q$ is n, the length of the $m-$sequence is $q^n - 1$.

The degree of $\underline{u}$ above is 4, because the length is $2^4 - 1 = 15$. We need a base sequence with a length $2^{n_1} - 1$ that divides $2^n - 1$, which we know happens when $n_1 | n$.

Covering Arrays

From Moura, Mullen, & Panario (2015):

Definition

A *covering array* $\mathrm{CA}_\lambda(N; t, k, s)$ is an $N \times k$ array on s symbols such that in every $N \times t$ subarray, each t-tuple of symbols occurs in at least λ rows. Then t is the *strength*, k the number of factors, λ the index and s the number of symbols.

Example of a covering array and application

An exhaustive testing of the factors shown below would be $2^4 = 16$ tests, but by following this covering array, all pair-wise interactions can be tested in just 5 tests.

	Web Server	Database	JDBC Driver	Directory Server
Test 1	IBM HTTP	DB2	IBM	Tivoli
Test 2	IBM HTTP	Oracle	Oracle	MS Active Directory
Test 3	Microsoft IIS	DB2	Oracle	MS Active Directory
Test 4	Microsoft IIS	Oracle	IBM	MS Active Directory
Test 5	Microsoft IIS	Oracle	Oracle	Tivoli

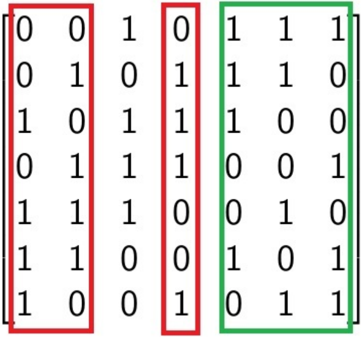
Creating covering arrays from $m-$sequences

m-sequences have nice properties that create covering arrays. This example, created from our previous example with length 7, has strength 2 everywhere, and strength 3 in certain columns.

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Which columns are covered?

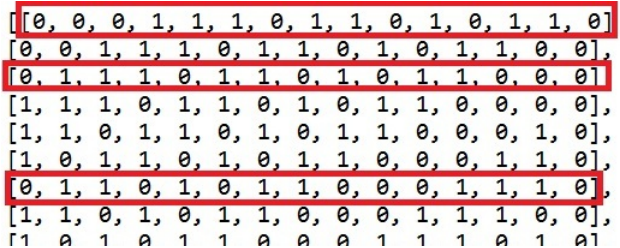
Munemasa (1998) shows that three columns are covered as long as they are not dependent. The original *m*-sequence was created with the primitive polynomial $x^3 + x + 1$. For example, any three consecutive columns (shown in green on the right) are covered. Any columns in the pattern $k, k+1, k+3 \pmod 7$ (shown in red on the left) are not covered.



$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Coverage of the interleaved sequence

This is the previous sequence, 0010111, interleaved with respect to the shift sequence $[0, 3]$, and turned into a covering array. As you might expect, if we choose only even or odd columns, we can look back to the original array to find out whether they are covered. In this diagram, we can see that columns 0, 2, and 6 are uncovered, corresponding to columns 0, 1, and 3 in the previous array.



```
[[0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0]
 [0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0],
 [0, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0]
 [1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0],
 [1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0],
 [1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0],
 [0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0],
 [1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0],
 [1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0]
```

Direction of research

From the properties of the base sequence and the shift sequence, how can we determine whether any set of columns in the interleaved sequence are covered or not?

References

Games, R. A: Cross-Correlation of M-Sequences and GMW-Sequences with the Same Primitive Polynomial. Discrete Applied Mathematics 12 (1985), 139-146.

Golomb S. W., Gong G.: Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar. Cambridge University Press, Cambridge (2005).

Gong, G.: Theory and Applications of *q*-ary Interleaved Sequences. IEEE Transactions on Information Theory, 41 (1995), 400-411.

He, J. J, Panario, D., Wang, Q.: A Family of Binary Sequences from Interleaved Construction and their Cryptographic Properties. Contemporary Mathematics 518 (2010), 209-223.

Munemasa, A: Orthogonal arrays, primitive trinomials, and shift-register sequences. Finite Fields and Applications 4(3), (1998), 252-260.

Thank You!