New Self-Shrinking Generator

Ali Adel Kanso
akanso@hotmail.com
Department of Mathematics, HCC
King Fahd University of Petroleum and Minerals
Hail, P. O. Box 2440, Saudi Arabia

Abstract

A new construction of a pseudorandom generator based on a single linear feedback shift register is investigated. The construction is related to the so-called self-shrinking generator. It has attractive properties such as conceptual simplicity, exponential period, and exponential linear complexity. The lower bounds are provided for the appearance of all patterns of reasonable length, and for some correlation attacks. The output sequences of this construction may have some applications in cryptography and spread spectrum communications.

Keywords: Stream ciphers, linear feedback shift registers, self-shrinking generator, and shrinking generator.

1 Introduction

In [1] a pseudorandom generator, the so-called self-shrinking generator, has been introduced by Meier and Staffelbach for potential use in stream cipher applications. The self-shrinking generator is attractive by its conceptual simplicity as it is based on a single linear feedback shift register (LFSR) [2]. Let LFSR $A$ be the linear feedback shift register defining the self-shrinking generator. Let $(A_t)$ be the original output sequence of LFSR $A$. The output sequence of the self-shrinking generator is generated by shrinking the original sequence $(A_t)$ as follows: The sequence $(A_t)$ is considered as pairs of bits $(A_0, A_1)$, $(A_2, A_3)$, $(A_4, A_5)$, .... If a pair $(A_{2i}, A_{2i+1})$ equals the value $(1, 0)$ or $(1, 1)$, it is taken to produce the bit 0 or 1, respectively. On the other hand, if the pair is equal to $(0, 0)$ or $(0, 1)$, it will be discarded.

In this paper, a new self-shrinking generator (referred to as NSSG) is introduced. The NSSG is a sequence generator composed of one single linear feedback shift register, say LFSR $A$, whose output sequence is shrunk in a similar way as is done for the self-shrinking generator of Meier and Staffelbach.

For the new self-shrinking generator the original sequence $(A_t)$ of LFSR $A$ is considered as pairs of bits $(A_0, A_1)$, $(A_1, A_2)$, $(A_2, A_3)$, .... If a pair $(A_t, A_{t+1})$ equals the value $(1, 0)$ or $(1, 1)$, it is taken to produce the bit 0 or 1, respectively. On the other hand, if the pair is equal to $(0, 0)$ or $(0, 1)$, it will be discarded.

Let $(S_t) = S_0, S_1, S_2, ...$ denote the output sequence of the NSSG of component sequence $(A_t)$. The new self-shrinking generator can be implemented as a special case of the shrinking generator of Coppersmith et al [3]. Recall that the shrinking generator uses two LFSRs $A$ and $B$ as basic components. The output bits are produced by shrinking the output of $B$ under the control of $A$ as follows: The output bit of $B$ is taken if the current output of $A$ is 1 otherwise it is discarded. Let $(A_t) = A_0, A_1, A_2, A_3, ...$ be the output sequence of the $m$-stage LFSR $A$ with initial state $A_0 = A_0, A_1, ..., A_{m-1}$ and feedback polynomial $f(x)$ defining the new self-shrinking generator. According to the rules of the new self-shrinking generator, the sequence $(A_t) = A_0, A_1, A_2, A_3, ...$ defines the control sequence, and the sequence $A_1, A_2, A_3, ...$ defines the generating sequence being controlled. Both sequences can be produced by the original LFSR $A$ when loaded with the initial state $A_0 = A_0, A_1, ..., A_{m-1}$ in the control register and the initial state $A_1 = A_1, A_2, ..., A_m$ in the generating register since one is a shift by one place of the other sequence.

This implies that the new self-shrinking generator can be implemented as a shrinking generator with two linear feedback shift registers having same characteristic feedback polynomial.

But the converse does not hold since a shrinking generator implemented as a new self-shrinking generator must have control sequence a translate of the generating sequence and this is not the case for a general shrinking generator.
2 Properties of the Output Sequence \((S_t)\) of the NSSG

In this section we analyse some of the properties of LFSR-based new self-shrinking generator. Suppose that LFSR \(A\) is a primitive \(m\)-stage linear feedback shift register with initial state \(A_0\) and characteristic feedback polynomial \(f(x)\) [2]. Let \((A_t)\) denote the output sequence of \(A\). Then \((A_t)\) is an \(m\)-sequence of period \(M = (2^m - 1)\). Let \((S_t)\) be the output sequence of this NSSG.

In the following lemmas, the period and the linear complexity of the output sequence \((S_t)\) are established. Finally, it is shown that the output sequence of the NSSG has good statistical properties.

2.1 Period and Linear Complexity of \((S_t)\)

We prove exponential bounds on the period and linear complexity of sequences produced by the NSSG. In the case of the period this bound is tight; for the linear complexity there is a gap by a factor of 2 between the lower and upper bound.

The importance of long period is to avoid repetition of the sequence after a short period of times. An exponentially large linear complexity avoids one of the more generic attacks on pseudorandom sequences and/or stream ciphers. Any sequence of linear complexity exponential large linear complexity avoids one of the more generic attacks on pseudorandom sequences.

Assume \(x = 1\) has occurred exactly once. It is well known that in a full period of an \(m\)-sequence of period \((2^m - 1)\) each \(m\)-bit pattern appears exactly \(2^{m-1}\) times, and the pair \(00\) appears exactly \(2^{m-2}\) times. By the definition of the new self-shrinking rule, it follows that the period of the new self-shrunken sequence \((S_t)\) divides \(2^{m-1}\).

In the next lemma, we show that the period of \((S_t)\) is exactly \(2^{m-1}\).

**Lemma 1** The period \(P\) of a new self-shrunken sequence generated by a primitive \(m\)-stage LFSR is equal to \(2^{m-1}\).

**Proof:** Since the \(m\)-stage linear feedback shift register is chosen to produce an \(m\)-sequence, then every non-zero \(m\)-bit pattern appears exactly once in a full period \(M = (2^m - 1)\) of the original \(m\)-sequence. Hence, in a full period of the original sequence the \(m\)-bit pattern \(11\ldots1\) appears exactly once. It is well known that in a full period of an \(m\)-sequence of period \((2^m - 1)\) each of the pairs \(11, 10,\) and \(01\) appears exactly \(2^{m-2}\) times, and the pair \(00\) appears exactly \(2^{m-1}\) times. By the definition of the new self-shrinking rule, it follows that the period of the new self-shrunken sequence \((S_t)\) divides \(2^{m-1}\).

Therefore, the period of the new self-shrunken sequence \((S_t)\) is \(P = 2^{m-1}\).

**Definition 2** The linear complexity \(L\) of a periodic sequence \((S_t)\) is equal to the degree of its minimal polynomial. The minimal polynomial is defined as the characteristic feedback polynomial of the shortest LFSR that can generate the sequence \((S_t)\).

**Lemma 3** The linear complexity \(L\) of a new self-shrunken sequence generated by a primitive \(m\)-stage LFSR satisfies: \(L > 2^{m-2}\).

**Proof:** Let \(Q(x)\) denote the minimal polynomial of the new self-shrunken sequence \((S_t)\). From the previous lemma the period of a new self-shrunken sequence \((S_t)\) generated by a primitive \(m\)-stage LFSR is \(P = 2^{m-1}\). Hence over \(GF(2)\), \((x^P - 1)\) can be written as \((x^P - 1) = (x - 1)^L\). Thus, the condition \(Q(x)\) divides \((x^P - 1)\) implies that \(Q(x)\) is of the form \(Q(x) = (x - 1)^L\) where \(L\) is the linear complexity of the sequence \((S_t)\). We claim that \(L > 2^{m-2}\).

Assume \(L \leq 2^{m-2}\). Then \(Q(x) = (x - 1)^L\) would divide \((x - 1)^{2^{m-2}} = (x^{2^{m-2}} - 1)\), but then the period of \((S_t)\) is at most \(2^{m-2}\) [see 4] contradicting lemma 1.

Therefore, the linear complexity \(L\) of the new self-shrunken sequence \((S_t)\) satisfies: \(L > 2^{m-2}\).

2.2 The Statistical Properties of \((S_t)\)

Suppose that the original sequence \((A_t)\) is an \(m\)-sequence of period \(M = (2^m - 1)\). In this section, we count the exact appearance of the number of ones and zeroes in a full period \(P = 2^{m-1}\) of the new self-shrunken sequence \((S_t)\). We
also show that in a full period of \((S_i)\) any subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}, S_{i+\beta-1}\) of length \(\beta \leq (m-k-1)\) (i.e. \((\beta+k+1) \leq m\) ), where \(k\) is the total number of zeroes in the subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}\), occurs \(2^{m-(\beta+1)}\) times for \(k = 0\), and at least \(\sum_{i=0}^{\lambda} \binom{k-1+i}{i} 2^{\lambda-i}\) times otherwise, where \(\lambda = m - (\beta + k + 1)\).

Note that
\[
\binom{k-1+i}{i} = \frac{(k-1+i)!}{(k-1+i-i)!} = \frac{(k-1+i)!}{(k-1)!i!}.
\]

The appearance of ones and zeroes in a full period of the output sequence \((S_i)\):

Since the original sequence \((A_i)\) is an m-sequence of period \((2^m-1)\), then in a full period of \((A_i)\) the 2-bit patterns 11 and 10 appear exactly \(2^{m-2}\) times. By the definition of the new self-shrinking rule, it follows that the number of ones and zeroes in a full period of the new self-shrunken sequence \((S_i)\) is \(2^{m-2}\).

Thus, the generated sequence \((S_i)\) is balanced.

**Lemma 4** Let the original sequence \((A_i)\) be an m-sequence of period \((2^m-1)\). Let \((S_i)\) denote the new self-shrunken sequence generated by self-shrinking the sequence \((A_i)\).

In a full period of \((S_i)\) any subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}, S_{i+\beta-1}\) of length \(\beta \leq (m-k-1)\) (i.e. \((\beta+k+1) \leq m\) ), where \(k\) is the total number of zeroes in the subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}\), occurs:

\[
\begin{align*}
&2^{m-(\beta+1)} \text{ times, for } k = 0, \\
&\text{and} \\
&\text{at least } \sum_{i=0}^{\lambda} \binom{k-1+i}{i} 2^{\lambda-i} \text{ times, otherwise.}
\end{align*}
\]

Where \(\lambda = m - (\beta + k + 1)\).

**Proof:** The original sequence \((A_i)\) is an m-sequence of period \((2^m-1)\). Thus, in a full period of \((A_i)\) each non-zero subsequence of length \(h \leq m\) occurs \(2^{m-h}\) times, and the all-zero subsequence of length \(h < m\) occurs \(2^{m-h-1}\) times [2].

Suppose we want to determine a lower bound on the number of times any subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}, S_{i+\beta-1}\) of length \(\beta\) occurs in a full period \(P = (2^m-1)\) of the new self-shrunken sequence \((S_i)\).

By the definition of the new self-shrinking rule, the generator produces an output bit whenever this bit is preceded by 1 in the original sequence \((A_i)\).

Let \(k\) be the total number of zeroes in the subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}\).

For \(k = 0\), the subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}, S_{i+\beta-1}\) (for \((\beta+1) \leq m\) ), in which \(S_i = S_{i+1} = \ldots = S_{i+\beta-2} = 1\), \(S_{i+\beta-1} = 1\) or \(S_i = S_{i+1} = \ldots = S_{i+\beta-2} = 1\), \(S_{i+\beta-1} = 0\), will occur in \((S)\) whenever the subsequence \(A_j, A_{j+1}, \ldots, A_{j+\beta-1}, A_{j+\beta}\), in which \(A_j = 1\), \(A_{j+1} = S_i\), \(A_{j+2} = S_{i+2}, \ldots, A_{j+(\beta-1)} = S_{i+(\beta-2)},\) and \(A_{j+\beta} = S_{i+(\beta-1)}\), occurs in the original sequence \((A_i)\). Obviously, this will occur \(2^{m-(\beta+1)}\) times in a full period of \((A_i)\). Hence, the subsequence \(S_i, S_{i+1}, \ldots, S_{i+\beta-2}, S_{i+\beta-1}\) will occur \(2^{m-(\beta+1)}\) times in a full period of \((S)\).
For \( k \neq 0 \), the subsequence \( S_j, S_{j+1}, \ldots, S_{j+\beta-2}, S_{j+\beta-1} \) (for \( (\beta + k + 1) \leq m \)) will occur in the new self-shrunken sequence \((S_j)\) at least whenever the subsequence \( A_j, A_{j+1}, \ldots, A_{j+(\beta+k-1)}, A_{j+(\beta+k)} \) (of length \( (\beta + k + 1) \)) less than or equal to \( m \), in which each bit of \( S_j, S_{j+1}, \ldots, S_{j+\beta-2}, S_{j+\beta-1} \) is proceeded by 1, occurs in a full period of the original sequence \((A_i)\). Moreover, the subsequence \( S_j, S_{j+1}, \ldots, S_{j+\beta-2}, S_{j+\beta-1} \) will occur from subsequences in which each 0 in \((A_i)\) is replaced by subsequences of 0's of length \( y \) where \( 1 \leq y \leq (\lambda + 1) \) and \( \lambda = m - (\beta + k + 1) \). [When doing that we have to make sure that the length of these subsequences of the original sequence \((A_i)\) does not exceed \( m \).

The total number of these subsequences in a full period of the original sequence \((A_i)\) is:

\[
\sum_{i=0}^{\lambda} \binom{k-1+i}{i} 2^{\lambda-i}.
\]

The subsequence \( S_j, S_{j+1}, \ldots, S_{j+\beta-2}, S_{j+\beta-1} \) may also occur from other subsequences such as \( A_j, A_{j+1}, \ldots, A_{j+(\beta+k-1)} \) (of length \( (\beta + k + 1) \)) for \( (\beta + k + 1) > m \).

Therefore, in a full period \( P = 2^{m-1} \) of the new self-shrunken sequence \((S_j)\) any subsequence \( S_j, S_{j+1}, \ldots, S_{j+\beta-2}, S_{j+\beta-1} \) of length \( \beta \leq (m-k-1) \) (i.e. \( (\beta + k + 1) \leq m \)), where \( k \) is the total number of zeroes in the subsequence \( S_j, S_{j+1}, \ldots, S_{j+\beta-2} \), occurs \( 2^{m-(\beta+1)} \) times for \( k = 0 \), and at least

\[
\sum_{i=0}^{\lambda} \binom{k-1+i}{i} 2^{\lambda-i}
\]
times otherwise, where \( \lambda = m - (\beta + k + 1) \).

### 3 Cryptanalysis

A suitable stream cipher should be resistant against a known-plaintext attack. In a known-plaintext attack the cryptanalyst can in some way detect a correlation between the known output sequence and the output of one individual LFSR. In order to assess the security of the generator we assume that the characteristic feedback polynomial of the LFSR is known. With this assumption we estimate the difficulty of finding the initial state of the LFSR.

In this section we discuss some approaches for possible cryptanalytic attacks and their complexities.

Assume that the original sequence \((A_i)\) is produced by a primitive \( m \)-stage LFSR (i.e. \((A_i)\) is an \( m \)-sequence of period \( 2^m - 1 \)). For cryptographic applications the key consists of the initial state and preferably the characteristic feedback polynomial of the LFSR. In order to assess the security of the generator we assume that the characteristic feedback polynomial is known. With this assumption we estimate the difficulty of finding the initial state of the LFSR.

We start with a general method for reconstructing the original sequence from the knowledge of a portion of the new self-shrunken sequence \((S_j)\).

Assume that \( S_0, S_1, \ldots, S_{n-2}, S_{n-1} \) is the known portion of \((S_j)\). The bit \( S_0 \) is produced by a bit pair \((A_{j_0}, A_{j_0+1})\) of the original sequence where the index \( j_0 \) is known.

Our aim is to reconstruct the original sequence in forward direction beginning with position \( j \). As we know \( S_0 \) we conclude that \( A_j = 1 \) and \( A_{j+1} = S_0 \). For the next bit pair \((A_{j+1}, A_{j+2})\) there remain one possibility if \( A_{j+1} \neq S_0 = 1 \) that is \( A_{j+2} = S_1 \), otherwise (if \( A_{j+1} = S_0 = 0 \)) there remain two possibilities that is \( A_{j+2} = 0 \) or \( A_{j+2} = 1 \), and so on.

Let \( k \) be the total number of zeroes in the subsequence \( S_0, S_1, \ldots, S_{n-2} \).

For \( k \neq 0 \), it can be shown by induction on \( k \) that in order to reconstruct the subsequence \( S_0, \ldots, S_{n-2}, S_{n-1} \) we have a total of:

\[
\Psi(m, k) = (m-2)^{k-1}(m+k-2)
\]
possible solutions.

For $k = 0$, to reconstruct the subsequence $S_0, S_1, ..., S_{n-2}, S_{n-1}$ we have only one possible solution (i.e. $\Psi(m, k) = 1$).

If a cryptanalyst obtains $m$ consecutive bits of the new self-shrunk sequence $(S_t)$ then, as $(S_t)$ is balanced, approximately half of these consecutive bits will be 0’s (i.e. $k = \frac{m}{2}$).

So in order to reconstruct these $m$ consecutive bits we have approximately a total of:

$$\Psi(m, \frac{m}{2}) = (m - 2)^{\frac{m-1}{2}}(m + \frac{m}{2} - 2)$$

possible solutions.

For security reason it is suggested to consider characteristic feedback polynomials of high hamming weight [7]. If the characteristic feedback polynomial is considered as part of the secret key, the reconstruction of the initial state has to be combined with an exhaustive search over all primitive characteristic feedback polynomials of degree $m$. Therefore, the complexity of the attack is increased by the factor $\frac{\varphi(2^m - 1)}{m}$, which is for large $m$ may be approximated by $2^n$. Hence, the total complexity is: $2^n \Psi(m, k) = 2^n (m - 2)^{k-1}(m + k - 2)$.

Thus, for maximum security, the key of the NSSG should consist of the initial state and the primitive characteristic feedback polynomial. Subject to these constraints the NSSG has a security level approximately equal to $2^n \Psi(m, k)$.

4 Related Work

Interesting example of existing LFSR-based constructions for comparison with the new self-shrinking generator is the self-shrinking generator of Meier and Staffelbach [2].

The advantage of the new self-shrinking generator over the self-shrinking generator is that, given an $m$-stage LFSR $A$ with initial state $A_0$ and primitive feedback polynomial $f(x)$ of degree $m$ that makes up a self-shrinking generator and a new self-shrinking generator. If the original sequence $(A_t)$ of $A$ has period $M = (2^m - 1)$, then for the self-shrinking generator in order to obtain a full period of the output sequence we have to clock LFSR $A$ $2M$ times, while for the new self-shrinking generator we only have to clock LFSR $A$ $M$ times. Also the output sequence of the self-shrinking generator has period $P$ that divides $2^{m-1}$ and satisfies $P \geq 2 \left\lceil \frac{m}{2} \right\rceil$, and linear complexity $L > 2 \left\lceil \frac{m}{2} \right\rceil - 1$, while the output sequence of the new self-shrinking generator has period $P = 2^{m-1}$ and linear complexity $L > 2^{m-2}$.

The disadvantage is that the self-shrinking generator is more secure against correlation attacks than the new self-shrinking generator [see 1].

5 Conclusion

From the theoretical results established, it is concluded that a NSSG with primitive LFSR generates sequences with large periods, high linear complexities, and good statistical properties. These characteristics and properties enhance its use in some applications in cryptography and spread spectrum communications.
References


