An Analysis of Linear Feedback Shift Registers in Stream Ciphers

Faheem Masoodi
Research Scholar
Department of Computer Science
A.M.U Aligarh

Shadab Alam
Research Scholar
Department of Computer Science
A.M.U Aligarh

M U Bokhari
Research Scholar
Department of Computer Science
A.M.U Aligarh

ABSTRACT
Linear Feedback Shift Registers (LFSRs) have always received considerable attention in cryptography. Owing to the good statistical properties, large period and low implementation costs, LFSR have achieved wide acceptance in developing stream ciphers. This paper intends to present a self-contained and comprehensive analysis of linear feedback shift registers and their application in stream ciphers. This research focuses on analyzing the mechanism of an LFSR, the two implementation variations and various properties of LFSR, which play a vital role in stream cipher design. In the last section of this paper, we address the security aspect of LFSR based stream ciphers and different techniques to enhance it.

General Terms
Design, Security, Cryptography, Algorithms

Keywords
LFSR, Linear complexity, period, NLFSR, Stream Cipher

1. INTRODUCTION
An LFSR is a shift register that, using feedback, elevates the bits through the register from current location to the next most-significant location, on each rising edge of the clock. Selected outputs (taps) are combined in an exclusive-OR (or exclusive-NOR) fashion to form a feedback mechanism, which causes the value in the shift register iterate endlessly through a sequence of unique values. An LFSR of any given size n (number of registers) is capable of producing every possible state during the period N=2n-1 excluding the all-zero state, such a sequence is called maximal sequence (abbreviated as m-sequence) [2].

Linear feedback shift registers as maximal length sequence generators are widely used in stream ciphers for key stream generation due to their good statistical properties, large period, low implementation costs, and are readily analysed using algebraic techniques. Maximal length sequences are generated when the LFSR passes through every non-zero state once and only once and are obtained when the feedback polynomial to which LFSR corresponds is primitive [1], a feedback polynomial of degree n is primitive if it is irreducible (cannot be factored) and has a period equivalent to 2n-1.

The predominant characteristic like large linear complexities, large period, statistical properties and pseudo randomness of the key stream generated by LFSR make LFSR a good choice for developing stream ciphers besides allowing a ready algebraic analysis of keystream generated by linear feedback shift registers

2. MECHANISM
The implementation of Linear feedback shift register consists of n input shift registers, where the input bit is calculated as a linear function of the content of the register. An n stage LFSR consists of clocked storage elements in the form of a shift register S and a feedback path in the form of tap sequence T where shift register $S=(s_n,s_{n-1},s_{n-2},\ldots\ldots..s_1)$ and a tap sequence $T=(t_n,t_{n-1},t_{n-2},\ldots\ldots..t_1)$, with each s_1 and t_1 being one binary digit. At each clock interval, all the bits are shifted right except bit s_1, which is appended to the key stream, and a new bit derived from S and T is fed back as input to the left end of the register. Whether a feedback is active or not is determined by feedback coefficients $T= (t_n, t_{n-1}..t_1)$

- If $t_i = 1$, feedback is active
- if $t_i = 0$, feedback is passive

A Linear feedback function produces a sequence S, satisfying the linear recurrence function. Assuming that the LFSR is initially loaded with some seed value $s_0,s_1,s_2\ldots s_{n-1}$. The next output bit is computed by the XOR-sum of products operation of storage elements and corresponding taps:

$$S_n \equiv s_{n-1}t_{n-1} + \cdots + s_1t_1 + s_0t_0 \bmod 2$$

Similarly next output is:

$$S_{n+1} \equiv s_nt_{n-1} + \cdots + s_2t_1 + s_1t_0 \bmod 2$$

Finally the general output can be shown as:

$$S_{n+1} \equiv \sum_{j=0}^{n-1} t_j \cdot S_{i+j} \bmod 2$$

Since the number of recurring states is finite, the generated sequence produced by LFSR must repeat itself after a finite period and also the length of the sequence is completely determined by the feedback coefficients and seed value

Theorem 1 *The maximum sequence length generated by an LFSR of degree n is* 2^n-1

Since an n-bit vector can assume only 2^n-1 states excluding an all-zero state, An n-bit LFSR can deterministically assume its

next state based on its previous state, as a result of which, as soon as an LFSR encounters a previous state, It starts to repeat itself. Therefore the maximum sequence length without repetition is 2^n-1. An all-zero state is discarded because if an LFSR assumes this state, it will get "stuck" and will never be able to leave this state.

An LFSR with a feedback coefficient vector $T(t_{m-1}\ldots t_1, t_0)$ is often specified with the help of polynomial:

$$P(x) = x^m + p_{m-1}x^{m-1} + \cdots p_1 x + p_0$$

In an attempt to achieve a maximal length sequence, the feedback polynomial of the LFSR should be primitive. Primitive polynomial is an irreducible polynomial of degree n, whose period is 2^n-1. The degree of the polynomial is the length of the shift register. An important aspect of irreducible and primitive polynomial is that all the primitive polynomials are irreducible but the reverse is not true. An irreducible polynomial p(x) of degree n is not said to be primitive if its order is not $2^n - 1$. [3]

Definition1. A polynomial of order *n*, having coefficients that are 0 or 1, is called 'irreducible' if it cannot be divided by another polynomial of degree *m*, where *m<n.* [4]

Definition2. An irreducible polynomial of order n, is called 'primitive', if and only if, it divides x^p+1 for only a *p* which is greater than or equal to 2^n-1 [4]

There may be more than one primitive polynomial of order n>2. The number of such primitive polynomials can be determined through the factorization theorem. The number of primitive polynomials of degree n is:

$$\frac{\emptyset(2^n - 1)}{n}$$

Where $\emptyset(x)$, known as the Euler function, denotes the number of positive integers less than the integer *x* and relatively prime to it. [6]

Modification of the feedback scheme allows us to implement LFSR in two different variations. Though the variations are cryptographically no better, but it can still affect the periodicity and software implementation [5]

Based on the configuration of gates and registers, LFSR can be divided in two categories.

- *The Fibonacci LFSR, also known as External-XOR LFSR or just LFSR*
- *The Galois LFSR, also known as Internal-XOR or canonical LFSR*

2.1 Fibonacci Functionality:

In the Fibonacci implementation, the taps are XORed sequentially with the output bit and the fed back into the leftmost bit. The shift register is initially loaded with bits a_0, $a_1\ldots$ a_{r-1} called the seed value (any value except all zeroes) and then clocked.

The output will be a pseudo random sequence and is given by the linear recurrence:

$$a_t = \sum_{i=1}^{r} q_i a_{t-i} \qquad \text{for } t \geq r.$$

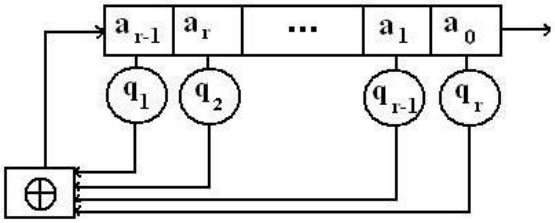

Fig 1: Fibonacci LFSR

2.2 Galois Functionality:

In the Galois implementation, with each clock cycle, bits that are not taps are shifted one position to the right unchanged. The taps on the other hand, are XORed with the output bit before they are stored in the next bit.

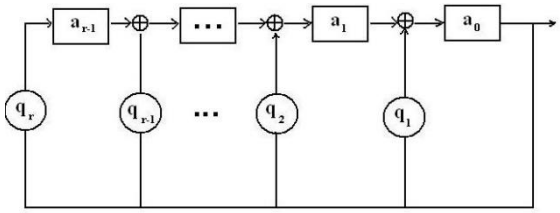

Fig 2 Galois configuration of LFSR

The shift register is initially loaded with bits a_0, $a_1\ldots$ a_{r-1} called the seed value (any value except all zeroes) and then clocked. If q_1, $q_2\ldots q_r$ are the feedback multipliers then the recurrence equations are as follows:

$$a_i' = a_{i+1} + q_{i+1}a_0, \qquad for\ 0 \leq i \leq r-2$$

$$a_{r-1}' = q_r a_0.$$

3. LFSR BASED STREAM CIPHERS:

An elegant way of realizing long pseudorandom sequences used by most but not all practical stream ciphers is to use linear feedback shift registers. LFSRs as maximal length sequence generators are commonly used as part of key stream generators in stream ciphers due to their good statistical properties, large periods and low implementation costs. Among stream ciphers that use LFSRs, we could cite the most famous case such as SNOW, SOSEMAUNK, GRAIN, A5/1, TRIVIUM, TURING and SOBER.

While designing cryptosystem it becomes imperative for cryptographers to consider suitable criteria for keystream generator to be used in cipher. Some of these design criteria are linear complexity, period and statistical measure of a keystream.

3.1 Linear complexity:

It is defined as length n of shortest LFSR that can mimic the generated output [4]. It is an indication for how difficult a sequence might be to replicate.

While a high linear complexity is a necessary condition, it is not a sufficient condition.

3.2 Period:

For an L stage LFSR, period is defined as the length of the stream, before it repeats itself. An LFSR with short Period results in encryption of different parts of plaintext with same keystream, which causes severe weakness. Practically, the period should be long enough to accommodate entire plaintext without repeating the keystream. The longest period possible corresponds to the largest possible state space, which is

produced by a maximal length tap sequence. Maximality of period guarantees good statistics.[6]

3.3 Statistical measures:

Once the sequence has been generated, it is required to assess it statistically, how well it is generated. Several statistical tests exist to determine the statistical behavior of the sequence. These include run test, rank test, frequency test, Fourier transform test, serial test, Lempel-Ziv complexity test and linear complexity test. [7] These tests generally check for random distribution, linear dependence among fixed length substrings, distribution of ones and zeroes in a sequence, the level of compression that can be carried out on tested sequence and whether a sequence is complex enough to be considered random

4. SECURITY ANALYSES OF LFSR BASED STREAM CIPHERS:

LFSRs are notoriously insecure from a cryptographic standpoint because the structure of an *n*-bit LFSR can be easily deduced by observing *2n* consecutive bits of its sequence using the Berlekamp-Massey algorithm [8]. Although Properties like large period, large linear complexity and a good statistical behavior are necessary but are not sufficient condition for a stream cipher to be considered cryptographically secure. Due to the inherent linearity, LFSR based stream ciphers are susceptible to several general attacks including known plaintext attack [9], algebraic attack [10], cache timing attack [11], fast correlation attack [18].

Cryptographically strong pseudo-random sequences are produced by combining more than one LFSR with some method to introduce non linearity. Three general methods of combining LFSRs employed to overcome the problem of linearity in LFSR based stream ciphers are: [13]

- Non linear combination generator
- Non linear filter generator
- Clock-controlled generator

4.1 Non Linear Combination Generator

The key stream is generated by manipulating the outputs of several parallel LFSRs using a non linear Boolean function *f*. The function *f* is called the combining function and maps one or more binary input variables to a binary output variable. The Boolean function must have a high algebraic degree, high nonlinearity and preferably a high order of correlation immunity. The keystream generated *z* is given by $z = f(x_1,x_2,\ldots x_n)$, where $x_1,x_2\ldots.x_n$ are the outputs of *n*-sub generators. General model for a nonlinear combination generator is shown in figure below

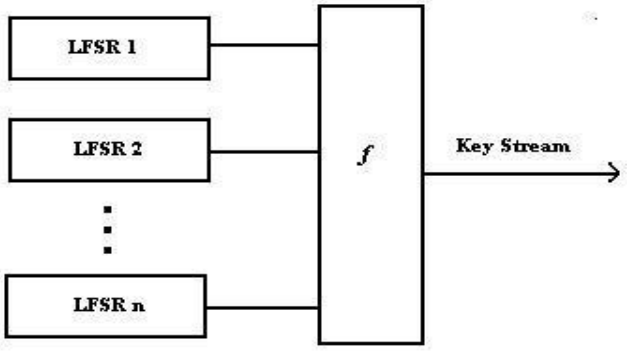

Fig 3: Non linear combination generator

If n maximum-length LFSRs with lengths $l_1,l_2,\ldots l_n$ are used together with the Boolean function f, the linear complexity of the keystream is

$$f(l_1,l_2,\ldots l_n) = a_0 + a_1l_1 + \ldots a_nl_n + \ldots a_{12\ldots n}l_1l_2\ldots..l_n$$

where $a_0,a_1\ldots a_n$ are the coefficients of in the algebraic normal form of *f*.[3]

4.2 Non Linear Filter Generator

A nonlinear filter generator uses a single maximum-length LFSR, and the keystream is generated as a nonlinear function f of the state of the LFSR.

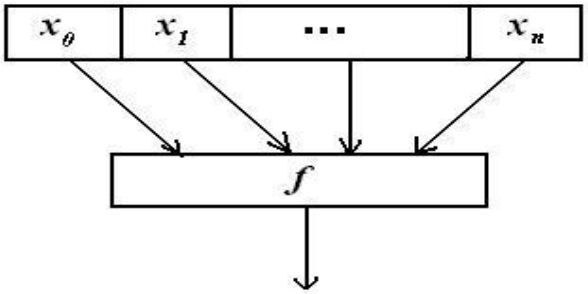

Fig 4: Non linear filter generator

The function f is called the filtering function. In this generator a single maximum length LFSR is used in contrast to the non linear combination generator where several LFSRs were used and different stages of single LFSR are used as input t the filtering function f. If a nonlinear filter generator is constructed by using a maximum-length LFSR of length L and a filtering function f of nonlinear order m then the linear complexity of the keystream is at most:

$$L_m = \sum_{i=1}^{m}\binom{n}{i} \quad [16].$$

4.3 Clock Controlled Generator

In clock controlled generators, the movement of data of one LFSR is controlled by the output of another LFSR. The register enabling clocking control is called as control register, and denoted as CR. The register which generates keystream according to the output sequence of CR is called as generator register, and denoted as GR. If *a(i)* represent the bit produced by CR and *b(i)* represent the bit produced by GR at instant *i*, the output of the keystream generator at time *i* is given as

$$u(i) = \sum_{k=1}^{i} a(k) \quad [1]$$

Since GR is clocked in an irregular manner, the output is the nonlinearly decimated sequence of a regularly clocked generator. General model for a nonlinear combination generator is shown in figure

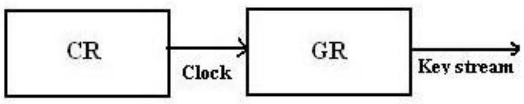

Fig 5: Clock controlled generator

5. CONCLUSION

Cryptographically secure pseudorandom number generators are of vital importance for stream cipher design and LFSRs though not secure due to their inherent linearity, are commonly used as part of key stream generators in stream ciphers due to their good statistical properties, large periods and low implementation costs. This study begins with the mechanism of an LFSR and its various classifications and goes into LFSR based stream ciphers, wherein we discuss the properties of LFSR like period, linear complexity and statistical behavior, which play a vital role in stream cipher design and we conclude by finally analyzing the security aspect of LFSR based stream ciphers and different techniques to introduce nonlinearity in the generated sequence, so as to make it cryptographically more secure

6. REFRENCES

[1] Sun Jing, Yang jing-yu, Fu De-sheng: Research On the Security of Key Generator in Stream Ciphers: The 1st International Conference on information Science and engineering (ICISE2009) pp. 1831—834

[2] P. P. Deepthi, Deepa Sara John and P. S. Sathidevi: Design and analysis of a highly secure stream cipher based on linear feedback shift register, Elsevier, computers and electrical engineering (2009), pp 235-243.

[3] Myat Su Mon Win: A New Approach to Feedback Shift Register: World Academy of Science, Engineering and Technology 48 2008 pp. 185—189

[4] A. Ahmad and A.M Elabdallai.: An Efficient Method to Determine Linear Feedback Connections in Shift Registers That Generate Maximal Length Pseudo-Random Up And Down Binary Sequences. Computer Electronic Engineering Vol.23, No.1 pp. 33-39, 1997

[5] Bruce Schneider: Applied Cryptography, Second Edition: Protocols, Algorithms, and Source Code in C. (1996)

[6] Kencheng Zeng, Chung-Hung Yang, Dah-Yea Wei and T.R.N Rao.: Pseudorandom Bit Generators in Stream-Cipher Cryptography: IEEE (1991)

[7] Cagigal, N.P; Bracho, S: Algorithmic Determination of linear feedback in a Shift Register for pseudorandom binary sequence generation: Electronic Circuits and Systems, IEE Proceedings. 1986 pp. 191 - 194 Vol. 133

[8] Elena Dubrova, A Transformation from the Fibonacci to the Galois NLFSRs, IEEE Transaction on Information Theory, Vol 55, No. 11, NOV 2009, pp 5263-5271

[9] C. Paar, J. Pelzl, Understanding Cryptography, Chapter 2-stream Cipher, Springer-Verlag, Berlin Heidelberg 2010 pp. 29-54

[10] Martin Voros: Algebraic Attack on Stream Ciphers, Master's thesis, submitted to COMENIUS UNIVERSITY, Department of Computer Science. 2007

[11] Gregor Leander, Erik Zenner, and Philip Hawkes: Cache Timing Analysis of LFSR-based Stream Ciphers, Proc. Crypto & Coding 2009, Springer LNCS 5921, 2009 .

[12] A. A. Bruen and R. A. Mollin, Cryptography and Shift Registers, The Open Mathematics Journal, 2009, pp 16-21.

[13] M U Bokhari and Faheem Masoodi.: Comparative Analysis of Structures and Attacks on Various Stream Cipher: Proceedings of the 4th National Conference; INDIACom-2010. pp. 236—238.

[14] Faheem Masoodi, Shadab Alam and M U Bokhari.: SOBER Family of Stream Ciphers: A Review: International Journal of Computer Applications (2011), Vol. 23.

[15] Mark Goresky: Fibonacci and Galois Representation of Feedback-with-carry shift registers, IEEE Transactions on Information Theory, Vol 48, Nov 2002

[16] Meltem Doğaner Özgan: A NEW NONLINEAR COMBINATION GENERATOR "MYBOUN", Master thesis, submitted to Graduate program in Electrical and Electronics Engineering, Bogazici University, 2006

[17] W. Liang and Long Jing, A cryptographic Algorithm Based on Linear Feedback Shift Register, 2010 International conference on computer application and system Modeling (ICCASM 2010), v15, pp 526-529

[18] Anne Canteaut, Fast correlation attacks against stream ciphers and related open problems. Theory and Practice in Information-Theoretic Security, 2005. IEEE Information Theory Workshop.