Varying Feedback Shift Registers

Yves ROGGEMAN
Université Libre de Bruxelles
Laboratoire d'Informatique Théorique
Campus Plaine - CP212
boulevard du Triomphe
B-1050 Bruxelles
Belgium

1. Context

It is well known that a stream cipher system can be described in terms of a Vernam scheme using a Pseudo-Random Number Generator as key generator. Each character m_t of the plaintext (viewed as an integer) is enciphered by adding the corresponding pseudo-random key character s_t. Deciphering is obtained by subtracting the same value stream from the ciphertext (see Fig.1).

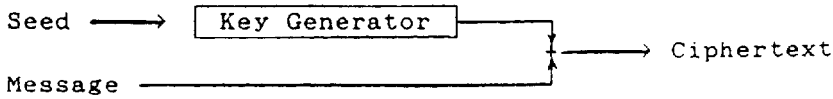

Fig.1

$$E(m_t) = m_t + s_t \qquad (1)$$
$$D(E(m_t)) = E(m_t) - s_t = m_t$$

A Lehmer Linear Congruential PRNG or a Linear Feedback Shift Register (LFSR) cannot be used in cryptographic systems because they can be cracked. In order to obtain Cryptographically Strong Number Generators, we can use Non-Linear Feedback Shift Registers. But a general model of such a NLFSR is difficult to implement and to study.

In another way, non-linearity is simulated in models involving more than one LFSR: product of sequences, cascade scheme, flip-flop, multiplexed LFSR, clock variation, a.s.o. But in most of these systems, every component can be isolated and/or the pseudo-random sequence is not always produced at a constant rate.

In this paper, we describe a new model based on FSR producing non-linear sequences, but which is easy to implement and can be used at a constant rate.

J.J. Quisquater and J. Vandewalle (Eds.): Advances in Cryptology - EUROCRYPT '89, LNCS 434, pp. 670-679, 1990.

2. Classical theory

2.1. FSR

A k-satge FSR is a machine involving k memory cells X_0, X_1, ... X_{k-1} (see Fig.2). At each clock pulse, every value is shifted one position left, the leftmost value is output and the rightmost cells is filled with a value depending on the k previous ones.

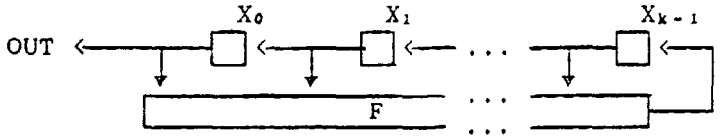

Fig.2

A solution (s_t) is an infinite sequence satifying

$$s_{t+k} = F(s_t, s_{t+1}, \ldots s_{t+k-1}) \qquad (2)$$

for some feedback function F. Such a solution is univokely determined by its initial state $[s_0, s_1, \ldots, s_{k-1}]$.

Classically, s_t belongs to a finite field GF(q) (GF(2) in most cases) and F is a rational function on GF(q). Such a register is noted $FSR^k(q)$.

2.2. Period and singularity

Each solution of a FSR is ultimately periodic. Its period and its singularity are the smallest integers π and σ satisfying

$$s_{t+\pi} = s_t \qquad \forall\ t \geq \sigma \geq 0,\ \pi > 0 \qquad (3)$$

A FSR is said to be non-singular if every solution is non-singular. This is achieved iff

$$F(\alpha, x_1, \ldots x_{k-1}) = F(\beta, x_1, \ldots x_{k-1}) \Longrightarrow \alpha = \beta \qquad (4)$$

In cryptographic applications, registers have to be non-singular and with maximal period.

2.3. Linear FSR

A k-stage Linear Feedback Shift Register $LFSR^k(q)$ is defined by

$$s_{t+k} = \sum_{i=0}^{k-1} c_i\ s_{t+i} \qquad t \geq 0,\ c_i \in GF(q) \qquad (5)$$

$$\text{or} \quad \sum_{i=0}^{k} c_i\ s_{t+i} = 0 \qquad \text{with } c_k = -1 \qquad (6)$$

Such a register is non-singular iff

$$c_0 \neq 0 \qquad (7)$$

The monic polynomial

$$f(x) = -\sum_{i=0}^{k} c_i\, x^i \qquad (8)$$

is called the characteristic polynomial associated with the $LFSR^k(q)$.

Its maximal period is (q^k-1) which is reached iff $f(x)$ is a so-called primitive polynomial on $GF(q)$.

The minimal polynomial of a periodic sequence is the characteristic polynomial of the smallest LFSR that can produced this sequence. Its degree is called the linear complexity of the sequence.

2.4. Transition matrix

The companion matrix C of $-f(x)$ is

$$C = \left[\begin{array}{c|ccc} 0 & & I & \\ \hline c_0 & c_1 & \ldots & c_{k-1} \end{array}\right] \qquad (9)$$

Its characteristic polynomial is $(-1)^k f(x)$, and its determinant is $(-1)^{k-1} c_0$. C is called the transition matrix of the $LFSR^k(q)$.

If we define the (transposed) state vector

$$\bar{s}_t' = [s_t, s_{t+1}, \ldots, s_{t+k-1}] \qquad (10)$$

the «'» indicating transposition, we have

$$\bar{s}_{t+1} = C\, \bar{s}_t \qquad (11)$$

The so-called generating functions for C^t and s_t are resp.

$$G(z) = \Sigma\, z^t\, C^t = (I - z\, C)^{-1} \qquad (12)$$

and

$$g(z) = \Sigma\, s_t\, z^t = \Phi(z)/\det(I-zC) \qquad (13)$$

where $\deg(\Phi) < k$.

3. Generalized LFSR

3.1. Non-degenerated solution

Generalizing Eq.11, we define a GLFSR^k(q) by

$$\bar{r}_{t+1} = M\ \bar{r}_t \qquad (14)$$

for any matrix M and $\bar{r}_t' = [r_{t,0}, \ldots r_{t,k-1}]$.

A GLFSR^k(q) is non-singular iff

$$\det(M) \neq 0 \qquad (15)$$

Let $R_0 = [\bar{r}_0, \bar{r}_1, \ldots \bar{r}_{k-1}]$ be the matrix whose columns are the first k states of a solution $(\bar{r}_t)$, this solution is called non-degenarated if $\det(R_0) \neq 0$.

We have the following property: if $(\bar{r}_t)$ is non-degenerated, it has the same minimal polynomial as M. Moreover, it is the characteristic polynomial of M.

3.2. Similar LFSR

Let C be the companion matrix of the monic characteristic polynomial of M. C is the transition matrix of a LFSR^k(q) similar to the GLFSR^k(q) defined by M. If $(\bar{r}_t)$ is a non-degenerated solution of M, let (s_t) be the solution of C corresponding to $\bar{s}_0' = [0, \ldots 0, 1]$ (the impulse), and let $S_0 = [\bar{s}_0, \ldots \bar{s}_{k-1}]$, we have

$$\begin{aligned} C &= S_0\ R_0^{-1}\ \ M\ \ R_0\ S_0^{-1} \\ \bar{r}_t &= R_0\ S_0^{-1}\ \ \bar{s}_t \end{aligned} \qquad (16)$$

Thus, each non-degenerated GLFSR^k(q) is similar to the LFSR^k(q) corresponding to the same characteristic polynomial.

3.3. Affine LFSR

Let A be any 1xk-matrix, $\bar{b}$ be any 1-vector and (s_t) be any solution of a LFSR^k(q). The 1-state sequence $(\bar{r}_t)$ defined by

$$\bar{r}_t = A\ \bar{s}_t + \bar{b} \qquad (17)$$

is the solution of a so-called Affine LFSR. It verifies

$$\sum_{i=0}^{k} c_i\ \bar{r}_{t+i} = \left(\sum_{i=0}^{k} c_i\right)\ \bar{b} = |c|\ \bar{b} = -f(1)\ \bar{b} \qquad (18)$$

If $R_t = [\bar{r}_t, \ldots \bar{r}_{t+k-1}]$ (a 1xk-matrix),

$$R_{t+1} = R_t\ C' \qquad (19)$$

4. Varying FSR

4.1. Definitions

Modifying Eq.2 as

$$s_{t+k} = F_t(s_t, s_{t+1}, \ldots s_{t+k-1}) \qquad (20)$$

we define a FSR with varying feedback functions (F_t). If there exist σ and τ such that

$$F_t = F_{t+\tau} \qquad \forall\ t \geq \sigma \qquad (21)$$

every solution of Eq.20 (s_t) is ultimately periodic and the register is called a Periodic FSR.

If $\sigma = 0$ and F_t is a linear function for every t, the PFSR is called a τ-PLFSR^k(q) defined by

$$\begin{array}{l} s_{t+k} = \Sigma\ c_{t,i}\ s_{t+i} = \overline{c_t} \cdot \overline{s_t} \\ \overline{c_t} = \overline{c_{t+\tau}} \qquad \forall\ t \end{array} \qquad (22)$$

Such a τ-PLFSR^k(q) is equivalent to a classical LFSR^k(q) iff $\tau = 1$.

4.2. Generating function

Let C_t be companion matrix of $\overline{c_t}$, the generating function associated with the τ-PLFSR^k(q) is

$$\begin{aligned} G(z) &= \Sigma\ z^t\ C_{t-1} \ldots C_1\ C_0 \\ &= (I + z\ C_0 + \ldots + z^{\tau-1}\ C_{\tau-2} \ldots C_0)\ (I - z^\tau\ C)^{-1} \end{aligned} \qquad (23)$$

where $C = C_{\tau-1} \ldots C_1\ C_0$.

Let $D_t = C_{t-1} \ldots C_1\ C_0$, we have $C = D_\tau$,

$$\overline{s_{t+1}} = C_t\ \overline{s_t} = D_{t+1}\ \overline{s_0} \qquad (24)$$

and $G(z) = \left(\sum_{t=0}^{\tau-1} z^t\ D_t\right)\ (I - z^\tau\ C)^{-1} = \Gamma(z)\ G_C(z^\tau)$ (25)

In Eq.25 we note $G_C(z)$ the generating function associated to the LFSR^k(q) with transition matrix C (see Eq.12).

If m(x) is the minimal polynomial of (s_t), and $m_C(x)$ the minimal polynomial of C,

$$m(x) \mid m_C(x^\tau) \qquad (26)$$

and the linear complexity of any solution of a τ-PLFSR^k(q) is at most τk

4.3. Period and singularity

The state of a τ-PLFSR$^k(q)$ does not only depend on the value of s_t. It includes the feedback index: t (mod τ). In this context, such a register is non-singular iff

$$c_{t,0} \neq 0 \qquad \forall\ t \qquad (27)$$

Let π be the actual period of any solution (s_t) of a τ-PLFSR$^k(q)$, its state-period is $\mu = \mathrm{lcm}(\pi, \tau)$.

It can be shown that any solution (s_t) of a τ-PLFSR$^k(q)$ with actual period π can be produced by a δ-PLFSR$^k(q)$ with $\delta = \gcd(\pi, \tau)$. Thus, if τ is a prime, any solution can either be produced by a classical LFSR$^k(q)$, or satisfies $\tau \mid \pi = \mu$.

5. Coupled LFSR

5.1. Definitions

In order to generate periodically varying feedback functions, we can use anaother LFSR. So we define a k,l-stage Coupled LFSR noted CLFSR$^{k,l}(q)$ as a τ-PLFSR$^l(q)$ where each c_t is the state of an Affine GLFSR$^k(q)$ (see Fig.3).

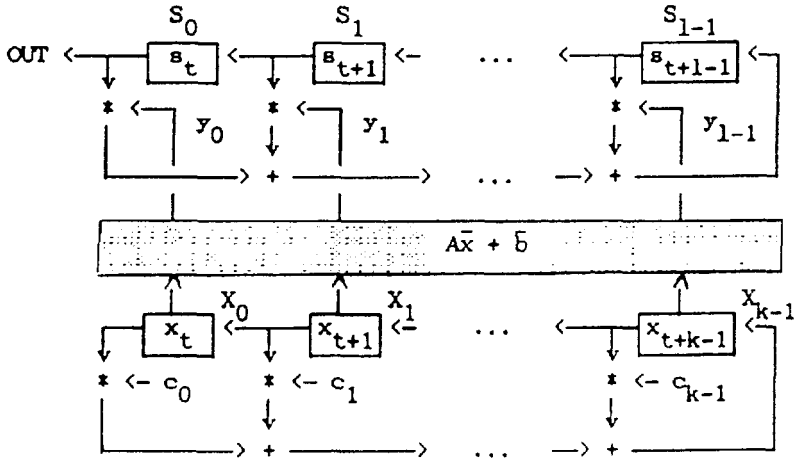

Fig.3

Such a model corresponds to equations

$$\begin{aligned} s_{t+l} &= \sum_{j=0}^{l-1} y_{t,j}\ s_{t+j} \\ y_{t,j} &= \sum_{i=0}^{k-1} a_{j,i}\ x_{t+i} + b_j \qquad j=0 \ldots l-1 \\ x_{t+k} &= \sum_{i=0}^{k-1} c_i\ x_{t+i} \end{aligned} \qquad (28)$$

It depends on kl+k+l parameters: the matrix A and the vectors $\overline{c}$ and $\overline{b}$.

Such a CLFSR^{k,l}(q) is non-singular iff

$$\begin{array}{lll} & y_{t,0} = b_0 \neq 0 & \forall\ t \\ \text{i.e.} & b_0 \neq 0 \text{ and } a_{0,i} = 0, & i=0\ldots k-1. \end{array} \qquad (29)$$

We shall only consider non-singular CLFSR.

5.2. Transition matrix

The state of a CLFSR^{k,l}(q) is given by the (k+l)-vector

$$\overline{v_t}' = [\ \overline{s_t}'\ |\ \overline{x_t}'\ ] = [s_t, \ \ldots\ s_{t+l-1}\,|\,x_t, \ \ldots\ x_{t+k-1}] \qquad (30)$$

So we define the associated transition matrix

$$T_t = \left[\begin{array}{c|c|c} 0 & I & \\ \hline b_0 & \overline{y_t}^{\cdot\prime} & 0 \\ \hline & 0 & C^{\cdot} \end{array}\right] \qquad (31)$$

where $\overline{y_t}^{\cdot\prime} = [y_{t,1}, \ \ldots\ y_{t,l-1}]$ and $C^{\cdot}$ is the classical transition matrix of the included LFSR^k(q).

If $\overline{b}^{\cdot\prime} = [b_1, \ \ldots\ b_{l-1}]$, we have

$$\begin{array}{l} \overline{y_t}^{\cdot} = A^{\cdot}\ \overline{x_t} + \overline{b}^{\cdot} \\ \overline{x_{t+1}} = C^{\cdot}\ \overline{x_t} \end{array} \qquad (32)$$

We now define the invertible matrices (33)

$$X_t = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & I & 0 \\ \hline \overline{x_t}^{\cdot} & 0 & I \end{array}\right], \quad Y_t = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline \overline{y_t}^{\cdot} & I & 0 \\ \hline 0 & 0 & I \end{array}\right]$$

$$A = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & I & A^{\cdot} \\ \hline 0 & 0 & I \end{array}\right], \quad B = \left[\begin{array}{c|c|c} b_0 & 0 & 0 \\ \hline \overline{b}^{\cdot} & I & 0 \\ \hline 0 & 0 & I \end{array}\right]$$

$$C = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & I & 0 \\ \hline 0 & 0 & C^{\cdot} \end{array}\right], \quad P = \left[\begin{array}{c|c|c} 0 & 1 & 0 \\ \hline I & 0 & 0 \\ \hline 0 & 0 & I \end{array}\right]$$

Using the commutator $[X_t, A^{-1}] = X_{t-1} A X_t A^{-1}$, we have

$$T_t' = C' Y_t P = C' B [X_t, A^{-1}] P \qquad (34)$$
$$= C' B [C^t X_0 C^{-t}, A^{-1}] P = C' B [X_0, C^{-t} A^{-1} C^t] P$$

In these formulae, we have the following properties:

- C' can be placed anywhere;

- B & $[X_t, A^{-1}]$ commute iff $b_0 = 1$ or $A^* = 0$;

- B & P commute iff $b_0 = 1$ and $\overline{b}^* = 0$ (i.e. $B = I$);

- P & $[X_t, A^{-1}]$ commute only if $\overline{x_t} \in \mathrm{Ker}(A^*)$.

This last property assure that if (x_t) is a solution of a primitive $\mathrm{LFSR}^k(q)$, and if $A^* = 0$, T'_{t+1} can never be expressed as a linear function of T'_t.

5.3. Statistical properties

In order to obtain the best statistical properties for the solution (s_t) of a $\mathrm{CLFSR}^{k,l}(q)$, we choose (x_t) as a solution of a primitive $\mathrm{LFSR}^k(q)$ which has period $\tau \square q^k - 1$.

For coupled registers, it can be proved that the period of any non-degenerated solution (s_t) is divisible by τ. The maximal period is then $(q^k-1)(q^l-1)$ which can be reached only if $(q-1) \mid l$.

There exist sufficient conditions to assure this maximal period, but they are not easy to verify. In practical applications however, $q = 2$ and we can choose k and l such that 2^k-1 and 2^l-1 are Mersenne primes. In this case, most solutions are maximal.

In a maximal solution, the distribution of multigrams $[s_t, \ldots s_{t+\mu}]$ satisfies:

- if $\mu < l$, the null multigram occurs $(q^{l-\mu-1}-1)(q^k-1)$ times
  and the other ones occur $q^{l-\mu-1}(q^k-1)$ times.

- if $\mu = l$, there exist

(q^l-q^{l-r}) multigrams occurring	$(q^{k-1}-1)$ times
$(q-1)(q^l-q^{l-r})$	q^{k-1}
$(q^{l-r}-1)$	(q^k-1)
$(q-1)(q^{l-r}-1)+q$	0

  where $0 \le r \le l-1$ is the rank of A^*. Bigger is r, more uniform is the multigrams distribution.

Moreover, the X^2 of the cross distribution of s_t and $s_{t+\mu}$ is

$$X^2(\beta) = \frac{(q-1)(q^l-1)}{(q^k-1)(q^l-q)^2}\,\{(q-1)(q^k-1)^2 - 2\,\beta\,(q^k-1)q^{l-2}(q^l-1) + \beta^2\,[((q^l-1)-(q-2))^2+(q-2)]\} \qquad (35)$$

where β eventually depends on the rank of $M = [\overline{b}^* | A^*]$.

If $\mu < l$, then $\beta = 0$ and $X^2(0)$ has the same constant value as in a classical primitive LFSR^l(q).

If $\mu = l$,
if rank(M) = r+1, then $\beta = 0$ as in the previous case;
if rank(M) = r, but $\overline{b}^* \neq 0$, then $\beta = q^{k-r}$;
if $\overline{b}^* = 0$, then $\beta = (q^{k-r}-1)$.

In the binary case, Eq.35 corresponds to the auto-correlation

$$P(\beta) = \frac{-1}{2^l-2}\,\left[1 - \beta\,\frac{2^l-1}{2^k-1}\right] \qquad (36)$$

Thus, the classical Golomb's theorems are locally satisfied for maximal solutions of a CLFSR.

6. Conclusion

Coupled Linear Feedback Shift Registers are simple designs involving LFSR producing non-linear pseudo-random sequences. They seem to be cryptographically strong enough for stream cipher systems, because the behaviour of one component is not independent from the other one. Nevertheless, a CLFSR can easy be implemented as a piece of hardware or software in a very efficient way.

7. Bibliography

GOLOMB, S.W., *Shift Register Sequences*, Holden-Day, San Fransisco, Calif., 1967.

MEYER, C.H. & TUCHMAN, W.L., "Pseudorandom Codes can be Cracked", *Elec. Des*, 23, Nov. 1972, p74-76.

REEDS, J.A., "«Cracking» a Random Number Generator", *Cryptologia*, 1, Jan. 1977, p20-26.

ROGGEMAN, Y., "Remarks on the Auto-Correlation Function of Binary Periodic Sequences", *Cryptologia*, 10 (2), April 1986, p96-100.

ROGGEMAN, Y., *Quelques classes de registres à décalage et leurs applications en cryptographie*, Ph. D. Thesis, Univ. Libre de Bruxelles, 1986.

RUEPPEL, R.A., *Analysis and Design of Stream Ciphers*, Springer Verlag, Berlin, 1986.