MATHEMATICAL APPROACH FOR RECOVERING ENCRYPTION KEY OF STREAM CIPHER SYSTEM

Abdul Kareem Murhij Radhi
College of Information Engineering, University of Nahrian ,Baghdad- Iraq.

Abstract

Stream cipher system plays an important role in many practical encryption systems. Moreover it can be an ideal source in random number generation. Shift registers are the backbone of such systems. This paper presents a method for attacking and recovering the basic key for the general stream cipher systems. Different criteria should be studied carefully when a key stream generated via these systems.

Depending on the type of feedback connection of different stages, the output stream may be classified as linear and nonlinear. Different parameters specify complexity degree of the output of these systems. Proposed system achieved to recover the initial cipher key via two modules. The First module focuses on simulating some general stream cipher systems including shift registers with different lengths, while the second recovers the basic or the initial key which is generated from the first module. Recovering key attached by normalizing polynomial equations to set of linear equations. Nonlinearity output normalized to linear equation before recovering the key. The proposed technique overcomes the complexity parameter of linear and nonlinear stream sequence compared with other techniques in this field. C++ version 4.5 where used in implementing the proposed system.

تقنية أسترجاع مفتاح التشفير لأنظمة التشفير المستمر بأستخدام الأتجاه الرياضي

عبد الكريم مرهج راضي
كلية هندسة المعلومات، جامعة النهرين. بغداد- العراق

الخلاصة

تلعب أنظمة التشفير المستمر دورا" مهما" في معظم أنظمة التشفير العملية. بالأضافة الى ذلك يمكن أن تكون تلك الأنظمة مصدرا" مثاليا" في أنظمة توليد الأرقام العشوائية. تعتبر مسجلات الأزاحة العمود الفقري لتلك الأنظمة. هذا البحث يقدم طريقة جديدة لمهاجمة وأستعادة المفتاح الأساسي لأنظمة التشفير المستمر العامة. يجب دراسة كيفية توليد مفتاح التشفير عند بناء تلك الأنظمة.

بالأعتماد على نوع الربط ين أجزاء مسجل الأزاحة ،فأن مخرجات هذا النظام يمكن أن تصنف الى خطية والى لاخطية. هناك عدة معايير تحدد درة تعقيد مخرجات هذا النظام. ينجز النظام المقترح في أستعادة المفتاح الأبتدائي من خلال جزئين أو وحدتين. الأول يركز على تمثيل بعض أنظمة التشفير المتضمنة مسجلات أزاحة مختلفة الأطوال، بينما الجزء الثاني يعمل على أستعادة المفتاح الأبتدائي الذي تم توليده في الجزء الأول. أن هذه الأستعادة تمت من خلال تبني طريقة رياضية جديدة وهي طريقة حل المعادلات الخطية بأستخدام طريقة كاوس

أو أيجاد معكوس المصفوفة. أن النظام المقترح قد تغلب على درجة التعقيد لأنظمة التشفير المستمر الخطية
واللاخطية مقارنة مع التقنيات الأخرى المستخدمة في هذا المجال. أستخدمت لغة C++ نسخة ٤,٥.

1. *Introduction*

LFSR provides a simple way to obtain sequences of vary high periods together with not vulnerable statistics properties. A set of *many* LFSR's as shown in (figure 1) is combined such that the output sequence in t_s seconds of LFSR from 0 to m_1 [٣] is

$$z(n) = s_1(n) \oplus s_2(n) \oplus .. \oplus {}^{s}\mathrm{m\_1}(n)..(1)$$ **where**

S_0(n), S_1(n) ... S_{M-1}(n) represent the outputs of LFSR 0 to M-1.

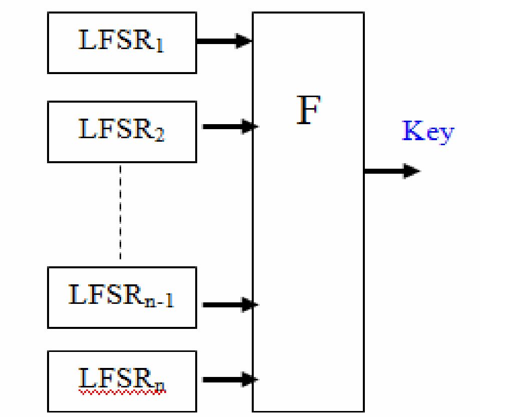

Figure 1: connected LFSRs by a function (F)

On the other hand, a stream cipher takes key K and initialization vector (IV) or V to produce the initial state as key stream generator produces a long output sequence from the internal state [1].

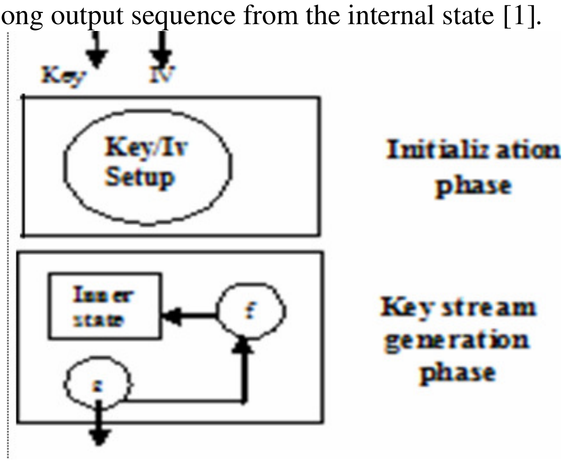

Figure 2:Two Phase key generator

٢.Representation of linear feedback shift register:

Feedback function in linear shift register can be written in the form f($s_0, s_{1,...,}s_{n-1}$) = $c_0s_0 + c_1s_1 + \cdots + c_{n-1}s_{n-1}$where each c_i is 0 or 1 and all addition is over GF(2).The constants c_0, c_1,.... c_{n-1} represent feedback coefficients. Linear feedback shift register can be represented by matrix structure as shown in the following form [3]:

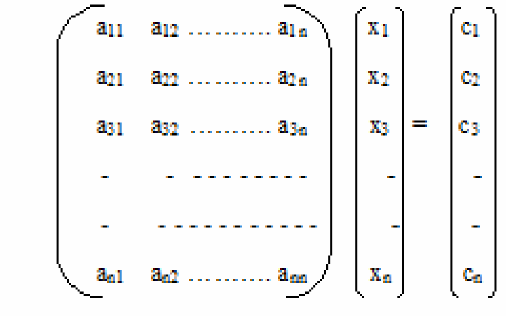

Figure ٣ :mathematical representation of LFSR

2.1 Generating Function with initial state:

In order to clarify representation of feedback function with initial state, consider initial conditions $a_0 = 1$, $a_1 = a_2 = a_3 = 0$ and taps $c_1 = c_4 = 1$, $c_2 = c_3 = 0$ so the feedback function is:

$$c_0s_0 + c_1s_1 + \cdots + c_{n-1}s_{n-1}$$

Generalizing Fibonacci recurrence an = an$_{\_1}$ + an$_{\_4}$ for n ~ 4[2], with binary stream ciphers are often constructed using linear feedback shift registers (LFSRs) since they can be easily implemented in hardware and can be readily analyzed mathematically. , however using LFSRs on their own is insufficient to provide good security. Therefore various schemes have been proposed to increase the security of LFSRs [3].

3.Proposed Technique

Proposed system aims to recover the initial key for key stream via the following two modules.

3.1-Simulation Module

(Figure 4) depict first module which represents how to simulate shift register structure, including its stages connection and its output via normalizing polynomial equations to first order equations, simplifying with following steps:

I. Identify LFSR length or number of shift register stages.
II. State linear feedback equation.
III. Specify output and number of rounds.

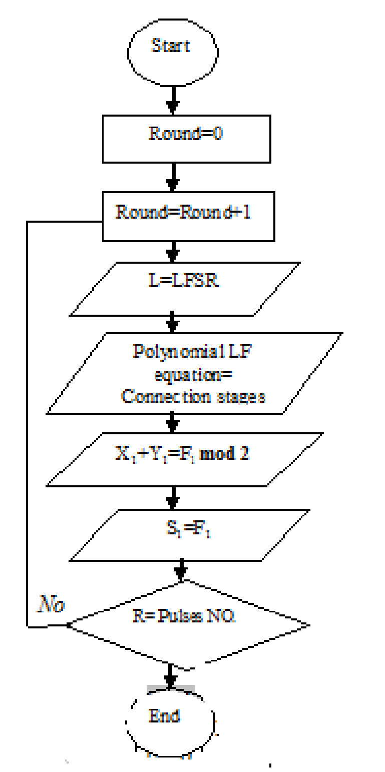

Figure 4:Simulation Mmodule1

3.2 Key Recovering Module

The second module *recovers initial state key* achieving the output of the previous module as follows:

I. Simulate the output of simulation module equivalent to shift register length. This simulation will be achieved via square matrix such that its dimension equivalents to shift register length.
II. Normalize and solve system of polynomial equations in step I by evaluating matrix inverse or Gauss elimination method for several equations.

٣,٣*Testing Examples*

The proposed technique tested through several examples. This paper will offer samples of them, as follows:

3.3.1 Example I

One of them if the LFSR of length five and feedback connection between stages two and five and the output from first stage: Then the output of five rounds are: (1 1 1 0 1). Then the Initial key is (1 0 1 0 1).

Table 1: Successive states of the LFSR With feedback coefficients (c_1, c_2, c_3, c_4, c_5) is (0, 1, 0, 0, 1, 0) and the Initial key is (1 0 1 0 1).

t	1	2	3	4	5	6	7	8	9	10	11
S_t	1	1	1	1	1	0	0	0	0	1	0
S_{t+1}	1	1	0	0	0	0	1	0	0	0	1
S_{t+2}	1	1	0	0	0	0	1	0	0	0	1
S_{t+3}	1	1	0	0	0	0	1	0	0	0	1
S_{t+4}	1	0	1	0	1	1	1	1	1	1	0

3.3.2 Example II

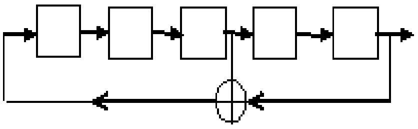

Figure5: LFSR with feedback coefficient(c_1,c_2,c_3,c_4,c_5)=(0,1,0,1,0) and binary initial state (1,0,1,1,1).

(Figure 5). presents LFSR with five stages: the proposed systems simulate the output feedback as:

$$S_{t+L}=C_{i+1}S_{t+2}+ C_{i+2}S_{t+4} \text{ Mod } 2 \ \forall \ t \geq 1 \ldots (3)$$

Where L represents shift register length which equal to five and t represent time pulse.

Table (1). offer successive states of the LFSR with previous features, while (figure 6). presents matrix representation of polynomial equations.

$$\begin{pmatrix} X_1 \oplus 0 \oplus 0 \oplus 0 \oplus 0 = 1 \\ 0 \oplus X_2 \oplus 0 \oplus 0 \oplus X_5 = 1 \\ X_1 \oplus 0 \oplus 0 \oplus X_4 \oplus 0 = 1 \\ 0 \oplus X_2 \oplus X_3 \oplus 0 \oplus X_5 = 0 \\ X_1 \oplus X_2 \oplus 0 \oplus X_4 \oplus 0 = 1 \end{pmatrix}$$

Figure 6: Matrix representation of equations

One of the subroutines to Normalizing and solving sets of equations shown in (figure 6). can be done via the following subroutine:

```
$\forall$ columns j
  begin from row i
If $\exists$ pivot $\in$ column j
     $\forall$ rows do
   If |row k | > |row max(k + 1)| do
     If A(maxi,j $\neq$ i, j) do
     Swap rows i and maxi
Divide each entry $\in$ row i by A[i, j]
$\forall$ rows
  Subtract row i from row i+1
End
```

Final matrix form obtained by the previous subroutine shown in (figure 7).

$$A=\begin{pmatrix} a_1 & 1 & 0 & 0 & ----- & 0 & 0 \\ a_2 & 0 & 1 & 0 & ----- & 0 & 0 \\ a_3 & 1 & 0 & 1 & ------ & 0 & 0 \\ - & - & - & - - - - & & - & - \\ a_{l-1} & - & - & - & - - - - & & 1 \\ a_l & - & - & - - - - - - & & & 0 \end{pmatrix}$$

Figure 7: Matrix normalization form

3.3.3 Example III

(Figure 9) presents LFSR with length 7 and initial state (1 0 1 0 1 0 1).

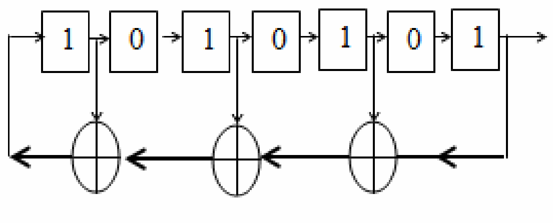

Figure 8: LFSR with length 7

$$\sum_{t\geq 0} S_t X^t = \frac{1+x+x^7}{1+x+x^3+x^3+x^7} \text{ --- (4)}$$

Since

$1+x+x^3+x^5+x^7=(1+x+x^7)(1+x^3+x^5)$

Then $\sum_{t\geq 0} S_t X^t =$

$$\frac{1+x+x^7}{(1+x+x^7)(1+x^3+x^5)} = \frac{1}{1+x^3+x^5}$$

This implies that (S_t) *is also* generated by LFSR with feedback polynomial $P_0(x) = 1+x^3+x^5$ as shown in (figure 9).

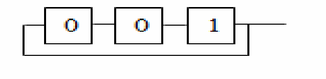

Figure 9: LFSR with length 3 generate the same sequence as LFSR with feedback polynomial of LFSR with length [7].

After that we will solve linear equations of the last polynomial equation as follows:

$$\left(\begin{array}{ccccccccccc} 0 & \oplus & 0 & \oplus & c_3x_3 & \oplus & 0 & \oplus & c_5x_5 & = & 0 \\ 0 & \oplus & c_2x_2 & \oplus & 0 & \oplus & c_4x_4 & \oplus & 0 & = & 0 \\ c_1x_1 & \oplus & 0 & \oplus & c_3x_3 & \oplus & 0 & \oplus & 0 & = & 0 \\ 0 & \oplus & c_2x_2 & \oplus & c_3x_3 & \oplus & 0 & \oplus & c_5x_5 & = & 0 \\ c_1x_1 & \oplus & c_2x_2 & \oplus & 0 & \oplus & c_4x_4 & \oplus & 0 & = & 1 \end{array}\right)$$

Figure 10:Polynomial equations of LFSR in figure9.

Table 2: Some of successive states of the LFSR With feedback coefficients (c_1, c_2, c_3, c_4, c_5) is (1, 0, 1, 0, 1) and the Initial key is (1 0 1 0 1).

t	1	2	3	4	5
S_t	0	1	0	1	0
S_{t+1}	0	0	1	0	1
S_{t+2}	0	0	0	1	0
S_{t+3}	0	0	0	0	1
S_{t+4}	1	0	0	0	0
S_{t+5}	0	1	0	0	0
S_{t+6}	0	0	1	0	0
S_{t+7}	1	0	0	1	0
S_{t+8}	0	1	0	0	1
S_{t+9}	1	0	1	0	0
S_{t+10}	1	1	0	1	0
S_{t+11}	0	1	1	0	1
S_{t+12}	0	0	1	1	0
S_{t+13}	1	0	0	1	1
S_{t+14}	0	0	1	1	0
S_{t+15}	1	0	0	1	1
S_{t+16}	1	1	0	0	1
S_{t+17}	1	1	1	0	0
S_{t+18}	1	1	1	1	0

Where maximum period of this LFSR with above features will be: P = 2^5 =2^L-1=31
The initial state will be as follows:
X_1= 1, X_2= 0, X_3=1, X_4= 0, X_5 = 1

4. Results and Discussion

Since LFSR can operate on any finite field, Galois field with 2^L elements [GF (2^L)] is appropriate field for representing LFSR items in this technique. The elements of this field and the coefficients of recurrence relation occupy exactly LFSR items.

GF (2^L) can be represented as modulo 2 coefficients of all polynomials with degree less than L.

The LFSR is mathematically equivalent to w parallel bit wide shift registers over GF (2). Testing the acquired results achieved via running the simulated system with any specific key stream period (pulses) and with defined shift register length(number of stages).Taking in account any feedback stages connection. Adapting the proposed method with any linear complexity and with any above shift register features will give the initial state (key) or the basic key stream of the first pulse.

This technique was being applied on natural number space via deterministic modification with carry (mod 2).

This is new technique not adapted from another technique; it is originality achieved using recovering mathematical automaton for recovering encryption or cipher key. Summarizing advantages of the proposed technique as follows:

1. The capability for recovering cipher key without restrictions (i.e. limitations of shift register length or type of feedback connection).
2. Different parameters specify complexity degree of the output for different symmetric cipher systems. This technique has the capability to recovering this cipher key for these systems (linear or nonlinear) with less complexity.
3. Moreover this technique can involves specific mathematical concepts to compress the processed growth data and space overflow such as reverse polish notation.
4. Capability of recovering initial state (internal state) with any length as shown in table (2).

Table (3). presents comparison parameters between previous approaches and the proposed technique [9]. It is clear that processing speed of the proposed technique is very fast and complexity measure equal to (2^L) where L represent shift register length. It is clear that this processing speed will be increased with the increasing of shift register length. Moreover it did not depend on shift register initial state contents.

Table 3: Stream cipher systems comparison

Stream cipher	Creation Date	ATTACK		
		Internal State	Attack technique	complexity
RC4	1987	2064	Key Derivation	2^{33}
A5/1	1989	64	KPA	2^{3991}
WAKE	1993	8192	CPA&CCA	Vulnerable
Rabbit	2003	512	N/A	N/A
Trivium	2004	288	Brute Force	2^{135}
Proposed System	2010	any	Mathematical	2^L

5. Conclusions and Suggestions

Proposed method presents a technique for recovering initial key generated by cipher stream key systems with low complexity comparing with previous approaches. The proposed system tested via different cipher systems using Turbo C++ version 4.5. Some suggestions in the future can be present to overcome complexity degree and storage cost such as adapting representation technique of operations ($\wedge$,$\vee$,, $\neg$,$\oplus$) and operands for instance reverse polish notation.

References

1. Fischer,S. and Khazaei,S.**2008**.” Chosen IV Statistical Analysis for Key Recovery Attacks on Stream Ciphers, AfricaCrypt**,** Casablanca - June 11-14.

2. L'Ecuyer , P. **2000**. A new Class of Linear Feedback Shift Register Generators, Proceedings of the Winter Simulation Conference.
3. Fischer, S.;khazaei, S. and Meier, W. **2008**," *Chosen IV Statistical Analysis for Key Recovery .Attacks on Stream Ciphers"*, Africa Crypt Casablanca June 11-14.
4. Beker, H. and Piper, F. **1982** *Cipher Systems*, Henery Beker and Fred Piper and Northwood Publications**,**.
5. Robshaw, J.B.M. **1995**. Stream Ciphers, RSA Laboratories Technical Report TR.701 Version 2.0, July 25.**.**
6. Klapper, A. and Xu, J. **1998**. Algebraic Feedback Shift Registers, Dept. of Computer Science, 763H Anderson Hall, University of Kentucky, Lexington, KY, 40506-0046, klapper@cs.uky.edu. Project sponsored by the National Science Foundation under grant number NCR-9400762,**.**
7. Sidek,A. and Sha'Ameri, A. **2007**.comparison analysis of stream cipher -algorithms for digital communication, *Journal technology*, 46(D) Jun**:** 1–16 © University Technology Malaysia.
8. Sutner,K.**2008**.Feedback Shift RegistersCarnegie Mellon University, fall.
9. Wikipedia of stream cipher system.
10. Grcar, J. F. **2010**. How ordinary elimination became Gaussian elimination. *Historia Mathematica*, in press.