Journal of ELECTRICAL ENGINEERING, VOL. 51, NO. 12/s, 2000, 81–84

ENHANCED CRYPTANALYSIS OF A CLOCK–CONTROLLED RUNNING KEY GENERATOR

Milan Vojvoda *

One simple running key generator which combines the outputs of two asynchronously clocked LFSRs has been proposed in [15]. In this paper the period of the keystream and several theorems concerning the number of runs in a ml-sequence are proved. Conditions for passing the Golomb's randomness postulates are proposed. Results of applied statistical tests (FIPS 140-1, gap test, serial correlation test) are presented. Finally, a modification of the generator using substitution of FCSRs (feedback with carry shift registers) for LFSRs is studied.

K e y w o r d s: stream cipher, running key generator

2000 *Mathematics Subject Classification*: 94A60, 94A55, 68P25

1 INTRODUCTION

LFSRs are still common building blocks of running key generators for binary additional stream ciphers [9], [12], [13]. They are fast, easy to implement in hardware, and allow the usage of algebraic methods in cryptanalysis.

One of such generators (denote it as G) was presented and studied in [16]. The generator G consists of two asynchronously clocked (in a stop-and-go fashion) LFSRs $L1$ and $L2$. Assume the polynomials $c_1(x), c_2(x) \in GF(2)[X]$ associated to the registers $L1$, $L2$ are primitive. Let us denote $\tilde{a} = \tilde{a}_0, \tilde{a}_1, \ldots$, resp. $\tilde{b} = \tilde{b}_0, \tilde{b}_1, \ldots$ the binary sequence produced by clock-controlled (as used in generator G) register $L1$ and $L2$. Moreover, denote $a = a_0, a_1, \ldots$ and $b = b_0, b_1, \ldots$ the binary sequence produced by the regularly clocked register $L1$, $L2$, respectively.

Algorithm of the generator G:

1. Keystream bit production: $s_t = L1(t) \oplus L2(t) = \tilde{a}_t \oplus \tilde{b}_t$.
2. Next-state function: if $s_t = 1$, then $L1$ clocks, otherwise ($s_t = 0$) $L2$ clocks.

2 PRELIMINARIES

Example 1. Assume the following realization of the generator G: $c_1(x) = 1 + x + x^2$ and $c_2(x) = 1 + x + x^3$. Let us look at the changes of the registers $L1$ and $L2$ states during the keystream generation. (Output bits $\tilde{a}_t$, resp $\tilde{b}_t$ are the underlined bits of the $L1$, resp. $L2$ states. The state of a register that clocks at a given time t is bold typed. The underlined bits of bold typed states of register $L1$ or $L2$ form runs (either blocks B_i^a or B_i^b, or gaps G_i^a or G_i^b) of the sequences a or b.

Table 1. The generation of the keystream of the generator G

t	State of $L1$	State of $L2$	s_t	Runs of a	Runs of b
0	$\underline{0}1$	**$\underline{0}$01**	0		G_0^b
1	$\underline{0}1$	**$\underline{0}$11**	0		
2	**$\underline{0}$1**	$\underline{1}11$	1	G_0^a	
3	$\underline{1}1$	**$\underline{1}$11**	0		B_1^b
4	$\underline{1}1$	**$\underline{1}$10**	0		
5	$\underline{1}1$	**$\underline{1}$01**	0		
6	**$\underline{1}$1**	$\underline{0}10$	1	B_1^a	
7	**$\underline{1}$0**	$\underline{0}10$	1		
8	$\underline{0}1$	**$\underline{0}$10**	0		G_2^b
9	**$\underline{0}$1**	$\underline{1}00$	1	G_0^a	
10	$\underline{1}1$	**$\underline{1}$00**	0		B_3^b
11	**$\underline{1}$1**	$\underline{0}01$	1	B_1^a	
12	**$\underline{1}$0**	$\underline{0}01$	1		
13	$\underline{0}1$	**$\underline{0}$01**	0		

Observation 2. *The keystream production could be characterized as joining of transformed runs of sequences a and b (look at the relation between underlined bold typed bits, s_t, and runs of the sequences a and b).*

Definition 3. If a polynomial $c(x) \in GF(2)[X]$, $\deg c(x) = |L|$ associated to a register L is primitive, then L is called a maximum-length (ml) LFSR. The output of an ml-LFSR with non-zero initial state is called an ml-sequence.

The following theorem characterizes the distribution of patterns in an ml-sequence.

* Department of Mathematics, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia, E-mail: vojvoda@kmat.elf.stuba.sk.

This research was supported by VEGA-grant 1/7611/20.

ISSN 1335-3632

Theorem 4. [11, p. 197, Fact 6.14] *Let u be an ml-sequence generated by a ml-LFSR L. Let k be an integer, $1 \leq k \leq |L|$, and let $\bar{u}$ be any subsequence of u of length $2^{|L|}+k-2$. Then each non-zero sequence of length k appears exactly $2^{|L|-k}$ times as a subsequence of u. Furthermore, the zero sequence of length k appears exactly $2^{|L|-k}-1$ times as a subsequence of u.*

3 RESULTS

Theorem 5. *Let u be an ml-sequence generated by an ml-LFSR with associated primitive polynomial $c(x)$, $\deg c(x) > 1$. Let $u_0 = u_1 = \cdots = u_{\deg c(x)-2} = 0$, $u_{\deg c(x)-1} = 1$. Then the number of runs in one period of the sequence u is even. Moreover, the number of blocks is equal to the number of gaps.*

P r o o f . This theorem follows immediately due to $u_{2^{\deg c(x)}-2} = 1$.

We determine the exact number of runs in one period of an ml-sequence in the following.

Theorem 6. *Let u denote a sequence generated by an ml-LFSR with associated primitive polynomial $c(x) \in GF(2)[X]$, $\deg c(x) > 1$. Assume $u_0 = 0$, $u_1 = 0$, ..., $u_{\deg c(x)-2} = 0$, $u_{\deg c(x)-1} = 1$. Then the exact number of runs in one period of the sequence u is $2^{\deg c(x)-1}$.*

P r o o f . According to Theorem 5, it is sufficient to prove that the number of blocks in one period of $u = u_0, u_1, \ldots, u_{2^{\deg c(x)}-2}$ or $u = u_0, u_1, \ldots, u_{2^{\deg c(x)}-2+\deg c(x)-1}$ equals $2^{\deg c(x)-2}$. Let $B^u[i]$ denote the number of blocks of length $\deg c(x) - i$. It follows from Theorem 4 that

$$B^u[i] = 2^{\deg c(x)-(\deg c(x)-i)} - \sum_{j=0}^{i-1} B^u[j](i-j+1)\,.$$

$$\sum_{j=0}^{i} B^u[j] = 2^{i-1}\,, \quad 1 \leq i < |L|\,, \quad B_0 = 1\,.$$

Corollary 7 (of Theorem 6)**.** *Let u and v be ml-sequences generated by ml-LFSRs L_u and L_v with associated primitive polynomials $c_u(x), c_v(x) \in GF(2)[X]$, $\deg c_u(x) = \deg c_v(x)$. Let $u_0 = u_1 = \cdots = u_{\deg c_u(x)-2} = 0$, $u_{\deg c_u(x)-1} = 1$ and $v_0 = v_1 = \cdots = v_{\deg c_v(x)-2} = 0$, $v_{\deg c_v(x)-1} = 1$. Then the sequences u and v have the same number of blocks, and gaps of lengths $1, 2, \ldots, \deg c_u(x) = \deg c_v(x)$.*

We generalize Theorem 6 for any non-zero initial state of the generating register.

Theorem 8. *Let u denote an ml-sequence generated by an ml-LFSR (from a non-zero initial state) with associated primitive polynomial $c(x) \in GF(2)[X]$, $\deg c(x) > 1$. Then the number of runs in one period of the sequence u is either $2^{\deg c(x)-1}$ or $2^{\deg c(x)-1}+1$.*

P r o o f . Let w denote a sequence u that starts with $u_0 = u_1 = \cdots = u_{\deg c(x)-2} = 0, u_{\deg c(x)-1} = 1$. Observe that any sequence u can be obtained from the sequence w by shifting [1, pp. 350–351]. The sequence w can be shifted to the beginning of a new run ($2^{\deg c(x)-1}$) or somewhere inside a run ($2^{\deg c(x)-1}+1$).

The following theorem concerning the period of the keystream of the generator G is based on Observation 2 and Theorem 8. (A conjecture was presented in [15].)

Theorem 9. *Assume that registers $L1$ and $L2$ with associated primitive polynomials $c_1(x)$ and $c_2(x)$, $\deg c_1(x), \deg c_2(x) > 1$ are loaded with a non-zero initial state. Then the period of the keystream sequence s of the generator G is*

$$(2^{\max\{\deg c_1(x),\deg c_2(x)\}}-1) + 2^{|\deg c_1(x)-\deg c_2(x)|}(2^{\min\{\deg c_1(x),\deg c_2(x)\}}-1)\,. \quad (1)$$

P r o o f . First, we prove that (1) is an integer multiple of the period of the keystream sequence s. There are 16 possibilities for the start and end runs of one period of the sequences a and b. Look at one of them (the other possibilities can be analyzed in a similar way).

Assume that a starts with a gap and ends with a block, b starts and ends with a gap. (The notation of blocks and gaps comes from Example 1.) $*$-denoted runs (as well as $\circ$-denoted, that have the start and end runs from different registers) clearly form an integer multiple of the period of the keystream sequence s.

Table 2. Joining of runs during the production of the keystream

$L1$		G_0^a		B_1^a	$\cdots$	$B^a_{2^{\deg c_1(x)-1}}$			G_0^a
$L2$	G_0^b		B_1^b		$\cdots$		$G^b_{2^{\deg c_2(x)-1}+1}$	G_0^b	
	$*$	$*$	$*$	$*$	$\cdots$	$*$	$*$		
		$\circ$	$\circ$	$\circ$	$\cdots$	$\circ$	$\circ$	$\circ$	

Finally, observe that the $\circ$-denoted part of the keystream sequence contains exactly one block of length $|L1|$ (if $|L1| \geq |L2|$) or one gap of length $|L2|$ (if $|L1| \leq |L2|$). Thus the $\circ$-denoted part of the keystream forms exactly one period.

The next theorem characterizes the basic balancedness of the keystream sequence.

Table 3. Results of statistical tests

#	Serial correlation test	Poker test (for quadruples)	Monobit test
0	0.002385	15.5264	9958
1	0.001986	6.08640	9734
2	-0.005419	9.92	10003
3	0.003574	10.8288	9750
4	0.007899	18.5472	9986
5	0.007897	16.6784	10156
6	0.004948	10.5088	10248
7	-0.002401	18.0352	10067
8	-0.000602	15.904	9973
9	-0.000602	16.6464	10224
	[-0.00068,0.00068]	[1.03,57.4]	[9654,10346]

Theorem 10. *Assume that registers $L1$ and $L2$ with associated primitive polynomials $c_1(x)$ and $c_2(x)$, $\deg c_1(x), \deg c_2(x) > 1$ are loaded with a non-zero initial state. Then the number of ones and zeros in one period of the generated keystream sequence is given as:*

1. $\deg c_1(x) \geq \deg c_2(x)$
   the number of ones: $2^{\deg c_1(x)} - 1$
   the number of zeros: $2^{|\deg c_1(x) - \deg c_2(x)|}(2^{\deg c_2(x)} - 1)$
2. $\deg c_1(x) < \deg c_2(x)$
   the number of ones: $2^{|\deg c_1(x) - \deg c_2(x)|}(2^{\deg c_1(x)} - 1)$
   the number of zeros: $2^{\deg c_2(x)} - 1$.

P r o o f . Theorem 10 follows from the proof of the period and from the production of the keystream as joining transformed runs from sequences a and b.

The following theorem about passing the first and second Golomb's postulates easily follows.

Theorem 11. *Assume that registers $L1$ and $L2$ with associated primitive polynomials $c_1(x)$ and $c_2(x)$, $\deg c_1(x), \deg c_2(x) > 1$, $|\deg c_1(x) - \deg c_2(x)| \leq 1$ are loaded with a non-zero initial state. Then the generated keystream sequence passes the first Golomb's postulate. Moreover, if $\deg c_1(x) = \deg c_2(x)$ then the keystream sequence passes the second Golomb's postulate, too.*

3.1 A note on the linear and sphere complexity

Theorem 12 [2, Theorem 3.4.4]**.** *Let N be an odd prime with $\gcd(N, q) = 1$, and let q be a primitive root modulo N. Then for any nonconstant sequence u of period N over $GF(2)$,*

1. *linear complexity is N or $N - 1$;*
2. *if $k < \min\{hwt(u_0, u_1, \ldots, u_{N-1}), N - hwt(u_0, u_1, \ldots, u_{N-1})\}$ then the sphere complexity of u is N or $N-1$, otherwise it is 0 (hwt denotes the Hamming weight).*

Theorem 12 puts important restrictions on the choice of the lengths of registers $L1$ and $L2$ (see Expression 1).

3.2 A note on the security of the generator

Assume now that the generator G consists of two subgenerators $G1$ and $G2$, respectively.

Using the known plaintext attack presented in [15] (another interesting attacks on stream ciphers are in [5], [14]) it is easy to find sequences a and b generated by these subgenerators $G1$ and $G2$. Thus the security of the whole generator against the known plaintext attack depends on the security of $G1$ and $G2$ against this kind of an attack.

Clearly, when using LFSRs $L1$ and $L2$ as the subgenerators $G1$ and $G2$, the key of the generator (the initial loading of the registers $L1$ and $L2$) is directly the beginning part of the sequences a and b.

If we use FCSRs (studied in [6], [7]) $F1$, $F2$ as the subgenerators $G1$, $G2$ the initial loading of the shift registers of $F1$, $F2$ can be obtained by the same way as described above. The remaining contents of the carry registers of $F1$, $F2$ can be found either by solving a set of equation or by exhaustive search.

4 STATISTICAL TESTS — RESULTS

The simulated realization of the generator G was: $c_1(x) = 1 + x + x^2 + x^5 + x^{19}$, $c_2(x) = 1 + x^3 + x^{31}$. The test set consisted of 1000 keystream sequences (each 20000 bits long) produced by this realization of the generator.

All of the tested sequences passed all tests given by FIPS 140-1 [3], 95% of them passed the serial correlation test [8] and none of them passed the gap test [8].

Table 3 outlines the values of the serial correlation coefficient, the statistics for the poker test [11, p. 182], and the number of ones in a sequence for the monobit test [3]. The last row of the table shows the expected intervals.

According to the results of the Maurer's universal statistical test [10] the keystream sequence is not significantly compressable (Q denotes the number of initial blocks, K denotes the number of tested blocks).

Theorem 13. *The keystream sequence produced by the generator G passes the long run test if $1 < |L1|, |L2| < 34$ (and registers $L1$ or $L2$ are loaded with a non-zero initial state).*

P r o o f . Observe that the longest run in the keystream sequence has $\max\{|L1|, |L2|\}$ bits (see Example 1).

5 CONCLUSION

In this paper we presented several theorems determining the number of runs in an ml-sequence. The period of the keystream sequence of the cryptanalysed generator is determined as well as its basic statistical properties. The results of statistical tests are outlined. The security of the generator against the known plaintext attack is generalized.

Table 4. Run test — numbers of occurrences of runs with certain lengths

# / run length	1	2	3	4	5	6
0	2459	1302	693	386	187	123
1	2642	1209	713	267	176	160
2	2340	1179	604	245	173	129
3	2400	1167	530	361	155	134
4	2589	1376	587	378	113	149
5	2512	1391	536	276	180	167
6	2680	1268	568	246	99	187
7	2540	1290	633	358	137	181
8	2397	1104	712	369	169	152
9	2645	1176	589	374	138	192
	[2267,2733]	[1079,1421]	[502,748]	[223,402]	[90,223]	[90,223]

Table 5. Maurer's universal statistical test - entropy on the 8-bit block

#	$Q = 2560$ $K = 256000$	$Q = 25600$ $K = 2560000$
0	8.003677	8.002048
1	7.999273	8.000793
2	8.002426	7.999964
3	8.000400	8.000941
4	8.001049	7.999997
5	7.999030	8.002516
6	7.998300	8.001611
7	8.000900	8.001506
8	7.999110	7.999905
9	8.002141	8.001281

Acknowledgement

The author would like to thank Professor Otokar Grošek for many helpful discussions.

References

[1] BIRKHOFF, G.—BARTEE, T. C.: Modern Applied Algebra, McGraw-Hill, New York, 1970. (in Slovak: Alfa 1981)

[2] CUSICK, T. W.—DING, C.—RENVALL, A.: Stream Ciphers and Number Theory, Elsevier Science B.V., 1998.

[3] FIPS 140–1, Federal Information Processing Standards Publication 140–1, NIST, 1994 (http://www.itl.nist.gov/fipspubs/fips140-1.htm).

[4] GROŠEK, O.—PORUBSKÝ, Š.: Cryptography — Algorithms, Methods, Practice (Šifrovanie — algoritmy, metódy, prax), Grada, Praha, 1992. (in Slovak)

[5] GROŠEK, O.: On Stability of Stream Ciphers (O stabilite prúdových šifier), Abstracts of the conference "Jesenný seminár z kryptoanalýzy" L. Mikuláš, October 9–11, 1996, pp. 14–26. (in Slovak)

[6] KLAPPER, A.: Feedback with Carry Shift Registers over Finite Fields, K.U. Leuven Workshop on Cryptographic Algorithms, Springer-Verlag 1995, pp. 170–178.

[7] KLAPPER, A.—GORESKY, M.: Large Period Nearly de Bruijn FCSR Sequences, Advances in Cryptology — EUROCRYPT'95 Proceedings Springer-Verlag 1995, pp. 263–273.

[8] KNUTH, D. E.: The Art of Computer Programming, vol. 2 Seminumerical Algorithms, Addison-Wesley, 1969.

[9] LIDL, R.—NIEDERREITER, H.: Introduction to Finite Fields and their Applications, Revised edition, Cambridge University Press, 1994.

[10] MAURER, U.: An Universal Statistical Test for Random Bit Generators, Advances in Cryptology — CRYPTO'90 Proceedings, Lecture Notes in Computer Science, Springer-Verlag 1991, pp. 409–420.

[11] MENEZES, A.—van OORSCHOT, P.—VANSTONE, S.: Handbook of Applied Cryptography, CRC Press, 1996, (www.cacr.math.uwaterloo.ca/hac).

[12] NEMOGA, K.: Linear Recurrent Sequences (Lineárne rekurentné postupnosti), Abstracts of the conference "Jesenný seminár z kryptoanalýzy" L. Mikuláš, October 9–11, 1996, pp. 1–13. (in Slovak)

[13] RUEPPEL, R. A.: Analysis and Design of Stream Ciphers, Springer-Verlag, 1986.

[14] SATKO, L.: A Correlative Attack of Siegenthaler and Rueppel (Korelačný útok Siegenthalera a Rueppela), Abstracts of the conference "Jesenný seminár z kryptoanalýzy" L. Mikuláš, October 9–11, 1996, pp. 27–35. (in Slovak)

[15] VOJVODA, M.: Cryptanalysis of a Clock-Controlled Running Key Generator, Journal of Electrical Engineering **50** No. 10s (1999), 16–18.

[16] VOJVODA, M.: Design and Cryptanalysis of a Stream Ciphers Generator (Návrh a kryptoanalýza prúdového šifrátora), Diploma Thesis, Bratislava, FEI-STU 1999. (in Slovak)

Received 3 July 2000

Milan Vojvoda is a graduate student of applied mathematics at the Faculty of Electrical Engineering and Information Technology of the Slovak University of Technology, Bratislava. His supervisor is Professor Otokar Grošek.