

A new linear consistency test attack on noised irregularly clocked linear feedback shift registers

Slobodan Petrović

NISlab, Department of Computer Science and Media Technology
Gjøvik University College, p.o. box 191, N-2802 Gjøvik, Norway
Email: slobodan.petrovic@hig.no

Abstract—Linear Consistency Test (LCT) is a widely used algebraic attack against pseudorandom generator schemes. A system of linear equations depending on a guessed part of the key is assigned to the analyzed generator and checked for consistency. If the guessed part of the key is not the right one, the system will be inconsistent with high probability. In the presence of noise, additional measures are necessary for this attack to be successful. They must reduce the influence of intercepted output bits complemented by noise. In this paper, a technique is described that tries to guess which bit(s) of the intercepted output sequence are complemented by noise and remove all the equations from the linear system assigned to the generator that depend on those bits. The technique is demonstrated on cryptanalysis of a Binary Rate Multiplier (BRM). The experiments on this generator show that such an attack is feasible if the noise level is up to moderate.

Index Terms—Cryptanalysis, Irregular Clocking, Linear Consistency Test (LCT), Linear Feedback Shift Register (LFSR)

I. INTRODUCTION

Linear Consistency test (LCT) is an algebraic attack against a pseudorandom generator scheme that tries to recover its whole initial state starting from some guessed bits of it. A linear system in the unknown bits of the key is set up, in which the right-hand side of every equation is an intercepted bit of the output sequence of the generator, and its consistency is checked. If the guessed part of the key is not the right one, such a system will be inconsistent with high probability, see [7]. The attack proceeds as follows: first, a subset of the key bits is guessed. Then, a system of equations, in which the rest of the key bits are variables is set up. If the guessed key bit subset is not the right one, the consistency probability of the obtained system will be very small. The consistency of the system is tested for all the possible choices of the guessed portion of the key. If the length of the intercepted output sequence of the generator is sufficient (see [7]), the right choice of the guessed part of the key will lead to a consistent system, whose solution will be the rest of the key.

If we consider a ciphertext-only attack scenario, in which some of the intercepted output bits of the generator are degraded (i.e. complemented) by noise, we have to compensate the influence of the complemented intercepted bits on the consistency of the system in order for the attack to be successful. In this paper, we first guess which of the intercepted bits are complemented by noise and then we remove from the system assigned to the generator all the

equations involving the guessed complemented bits. Suppose that the length of the intercepted output sequence is sufficient for making decisions about consistency of the system assigned to the analyzed generator. If the choice of the guessed portion of the key is right, and the guess of the positions of the intercepted bits affected by noise is right, the remaining system will be consistent. If the guessed portion of the key is the right one but the choice of the positions of the intercepted output bits degraded by noise is wrong, the resulting system might be consistent (the probability for this is significant). But if the guess of the key portion is wrong, the system will remain inconsistent with high probability even if the guess of the positions of the bits of the intercepted output sequence degraded by noise is right.

We apply the attack described in the previous paragraph on a representative of a special class of pseudorandom sequence generators, so-called *irregular clocking* generators. Specifically, we analyze how the new LCT attack can be applied on a noised Binary Rate Multiplier (BRM) [2], but the same ideas can be applied in attacks against other representatives of the class as well - the Stop/Go generator, the Shrinking generator, the Alternative Step generator and so on.

The attacks against BRM have been studied by many authors, since that scheme is widely used in practice due to the desirable properties of the output sequence achievable with it (extremely long period and linear complexity, good statistical properties etc.) Most attacks against BRM are *correlation attacks* (for example [3], [4], etc.) The LCT attack against BRM has also been attempted [5], but in a known-plaintext attack scenario, i.e. without noise. The possibility of an LCT attack against noised BRM was studied in [1] and [6]. This attack tries to avoid influence of the bits of the intercepted sequence affected by noise by changing the starting point in the intercepted sequence from which the setting up of the system of equations starts. The attack might be successful if the noise level (i.e. the probability of '1' in the noise sequence) is small, but false alarms and missing the event regarding the right guess of the part of the key are inevitable and because of that a more precise localisation of the bits affected by noise is needed.

The structure of this paper is as follows: In Section II, the BRM-based pseudorandom sequence generator is described. In Section III, the detailed description of the new attack is given. In Section IV, the experimental results are presented

and discussed. Section V concludes the paper.

II. THE ANALYZED GENERATOR

We analyze the LCT attack on a noised pseudorandom sequence generator involving a primitive known as The Binary Rate Multiplier (BRM). BRM consists of 2 linear feedback shift registers (LFSRs). One of them, the clocking LFSR (LFSR_s), determines the clocking sequence for the clocked LFSR (LFSR_u), see Fig. 1.

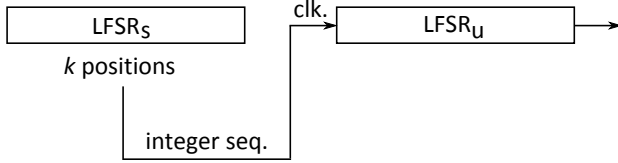


Fig. 1. The BRM primitive

The BRM operates as follows (Fig. 2): Without clocking by LFSR_s, the register LFSR_u produces the binary sequence u_n . At the clock pulse i of LFSR_s, the bits from k positions of LFSR_s determine the integer s_i that represents the number of bits from the sequence u_n that are going to be discarded. The integers s_i , $i = 1, 2, \dots$ make the sequence s_n . The process of discarding bits in this way is called *non-uniform decimation* of the sequence u_n . The maximum value of the integer s_i determines the maximum number of bits from the sequence u_n that can be discarded at a time. The binary sequence z_n is the output sequence of the whole BRM.

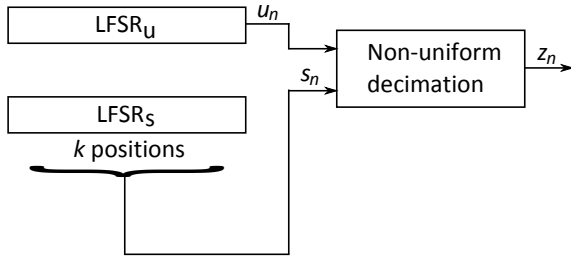


Fig. 2. Operation of the BRM

The BRM primitive has become popular in the design of stream ciphers since it can be shown [2] that the produced sequence z_n has extremely long period and high linear complexity preserving at the same time good statistical properties of a single LFSR.

III. THE NEW LCT ATTACK

In this section, we give details of the new LCT-based attack against a noised BRM. The general description and remarks about LCT have been exposed in the Introduction. To design an LCT attack against BRM, we have to determine which part of the BRM key (which consists of the initial states of LFSR_s and LFSR_u, as usual) is to be guessed. It is shown in [5] that assigning a linear system to a BRM when the initial state of LFSR_s is guessed is easy. Then the unknowns in the

system are the bits of the output sequence of LFSR_u without decimation together with the bits of the initial state of the same LFSR and the right-hand side of any equation in the system is the corresponding bit of the intercepted sequence. In our new LCT attack on a noised BRM, we use the same approach. We guess the initial state of LFSR_s and make a system of linear equations in the unknowns of the initial state of LFSR_u and the unknown bits of the output sequence of LFSR_u without decimation. The main point of our attack is the algorithm that eliminates the influence of the bits of the intercepted sequence complemented by noise.

Example 1

Suppose the BRM from Fig. 1 uses 4-bit LFSRs and the primitive feedback polynomials of LFSR_s and LFSR_u are $f_s(x) = 1 + x + x^4$ and $f_u(x) = 1 + x^3 + x^4$, respectively. Let the number of output taps of LFSR_s be $k = 2$ and the tap positions are the first and the second (from the left). Let the initial states of LFSR_s and LFSR_u be 1010 and 0110, respectively. Then the clocking sequence for LFSR_u (i.e. the integer sequence s_n) is 31021002333... and the output sequence of the BRM is 11010110111...

Let the cryptanalyst's guess of the initial state of LFSR_s be right, i.e. 1010. In the LCT attack against the generator without noise, the so-called *decimation sequence* is generated, containing the symbol '2' in the positions of the unknown bits. Each symbol '2' will correspond to a new variable in the system of equations assigned to the generator. In this case, the decimation sequence will be 2222 | 22212102212011220222122212221... The symbol | delimits the variables of the initial state of the clocked register LFSR_u from the rest of the variables. The variables to the left from the symbol | are given in the order x_4, x_3, x_2, x_1 , whereas the variables to the right from the symbol | are given in the increased order of indexing, i.e. x_5, x_6 , etc. Then the system of linear equations assigned to the given BRM is:

$$\begin{aligned} x_3 + x_4 + x_5 &= 0 \\ x_2 + x_3 + x_6 &= 0 \\ x_1 + x_2 + x_7 &= 0 \\ x_1 + x_5 &= 1 \\ x_5 + x_6 + x_8 &= 0 \\ &\vdots \end{aligned}$$

□

The new ciphertext-only attack against a BRM is described below:

1. Guess the initial state of the LFSR_s.
2. Set up a system of equations assigned to such a BRM without involving the intercepted bits. Such a system is homogeneous and always consistent.
3. Set up a system of equations involving only the equations containing the intercepted bits.
4. Join the obtained systems and check the consistency of the joint system. The following cases are possible:
 - 4.1 There is no noise and the right initial state of LFSR_s was guessed - the joint system will be consistent and

the missing bits of the initial state of LFSR_u can be obtained by solving the system.

- 4.2 There is no noise and the guess of the initial state of LFSR_s was wrong - the joint system will be inconsistent with high probability, see [7].
- 4.3 The intercepted sequence was degraded by noise and the guess of the initial state of LFSR_s was right - the joint system will be inconsistent with high probability because of the bits of the intercepted sequence complemented by noise.
- 4.4 The intercepted sequence was degraded by noise and the guess of the initial state of LFSR_s was wrong - the joint system will be inconsistent with high probability.
5. Suppose that t bits of the intercepted sequence were degraded by noise. Guess their positions. Remove all the equations involving these complemented bits from the joint system. Provided the intercepted sequence is long enough (see [7]), the following cases are possible:
 - 5.1 The guess of the initial state of LFSR_s was right - if the guess of the positions of the bits of the intercepted sequence that were complemented by noise was right, the system will become consistent and solving the system will give the missing bits of the key of the BRM. If the guess of the positions of the bits of the intercepted sequence complemented by noise was wrong, there will be relatively high probability (compared with that in the case of a wrong guess of the initial state of LFSR_s) that the system will become consistent.
 - 5.2 The guess of the initial state of LFSR_s was wrong - the system will not become consistent even if the guess of the positions of the bits of the intercepted sequence complemented by noise was right.
6. Repeat the step 5. of the algorithm for all the combinations of guesses for the positions of the t bits complemented by noise in the intercepted sequence, starting from $t = 1$, then $t = 2$ and so on, until a consistent system is obtained.

The success of the attack described above depends on the level of noise, i.e. the ratio between the number of bits complemented by noise in the intercepted sequence and the length of the intercepted sequence. A relatively small level of noise ensures relatively small number of combinations for the guesses of the complemented bits, which makes the attack feasible. In that case the attack is likely to be successful.

Example 2

Refer to the Example 1 above and suppose that the first bit of the intercepted output sequence from the generator is complemented by the noise sequence. This affects the equation $x_1 + x_5 = 1$ from Example 1 and other equations involving the complemented bit (2 more equations, since the weight of the feedback polynomial of LFSR_u is 2). The system of linear equations assigned to the given BRM becomes inconsistent (with high probability). To mitigate this, we have to remove the equations from the system that involve the complemented bits (in our example, the number of complemented bits in the

intercepted sequence is $t = 1$). We guess the position of the complemented bit. Suppose our guess is right, i.e. the first bit in the intercepted sequence is complemented by noise. If we remove all the equations involving that bit from the system, it will become consistent again, since, as we said in Example 1, our guess of the initial state of LFSR_s was right. By solving the system, we get the initial state of LFSR_u . If the guess of the position of the complemented bit in the intercepted sequence is wrong, there is some probability that the system remains consistent if the guess of the initial state of LFSR_s was right. \square

Regarding the time complexity of the attack, we should note that it is necessary to check all the possible initial states of LFSR_s in the attack, which gives the time complexity of the attack of $O(c \cdot 2^{L_s})$, where L_s is the length of LFSR_s and c is the number of combinations for complementing the bits of the intercepted sequence. c is small for low levels of noise, which makes the attack feasible in those cases.

IV. EXPERIMENTAL WORK

The experimental setup involved LFSR_s and LFSR_u , both of length 4, with primitive feedback polynomials ($f_s(x) = 1 + x + x^4$ and $f_u(x) = 1 + x^3 + x^4$). 2 positions of the register LFSR_s determined the clocking of LFSR_u , i.e. $k = 2$, which means that up to 3 bits of the output sequence of LFSR_u without decimation could be discarded at a time. The length of the intercepted output sequence was 20 and up to 2 bits of the output sequence of the BRM could be complemented by the noise, i.e. $p(1) \leq 10\%$ in the noise sequence. The experiment consisted of the following: for a fixed number of bits of the intercepted sequence complemented by noise t , $t = 1, 2$, all the possible combinations of positions of bits complemented by noise were tried. For each such combination, after complementing the intercepted sequence bits accordingly, the new LCT attack was run and the number of cases in which the guess of the initial state of LFSR_s was right and the consistent system was obtained was recorded. The number of cases in which the guess of the initial state of LFSR_s was wrong and a consistent system was obtained (false alarms) was also recorded. The cases where the number of false alarms was greater than the number of correct guesses followed by consistent systems were of particular interest, since in such a case the solution of the cryptanalytic problem given by our algorithm would be wrong. The goal of the experimental work was to investigate how the number of such cases behaves when t increases. In addition to the number of false alarms, the number of initial states of LFSR_s that the LCT-based attack algorithm would label as solution states would be important to study since in the case of a false alarm, the cryptanalyst would have to check those states further in order to eliminate the wrong solutions. We should bear in mind that the right solution is always offered by the attack algorithm, even when we get false alarms. The experimental results are given in Table I. In that table, for $t = 1$ and $t = 2$, the numbers of false alarms n_f are listed. In addition, the maximum numbers of solutions (i.e. the initial states of LFSR_s offered by the attack

TABLE I
THE NUMBERS OF FALSE ALARMS OBTAINED WITH THE NEW LCT
ATTACK (SEE TEXT)

	N	n_f	$n_f\%$	n_s	f_s	$f_s\%$
$t = 1$	20	1	5	2	1	5
$t = 2$	190	62	33	5	2	1

algorithm) n_s are given together with the numbers of cases f_s , in which the false alarms were generated with the maximum number of solutions (that would be the worst possible case for the cryptanalyst). In the table, N represents the total number of possible combinations for complementing the bits of the intercepted sequence.

From the Table I it can be noted that for $t = 2$ the number of false alarm cases is quite high, approx. 33% of all the cases. The maximum number of solution initial states of LFSR_s offered by the attack algorithm in that case is $n_s = 5$, which is 1/3 of the possible number of initial states of that register. This number is also quite high, but to recover from such a situation, ordering of these solution states according to the corresponding numbers of consistent systems obtained in the attack is possible to perform, which in most cases places the right initial state to the second position. This eliminates the problem of too many solutions offered by the attack algorithm and also makes the problem of too many false alarms easier. The greatest advantage of the new attack compared with the attack from [1] and [6] is in the fact that the new algorithm always produces the right solution, even when it is accompanied by a false alarm.

V. CONCLUSION

In this paper, a new Linear Consistency Test (LCT)-based attack against a noised pseudorandom generator scheme employing irregular clocking is described and analyzed. The attack was applied against a specific representative of this class of generators known as The Binary Rate Multiplier (BRM). The attack first assigns a system of linear equations to the BRM based on a guessed initial state of its clocking LFSR. This system is then checked for consistency and if consistent, the right initial state of the clocking LFSR was guessed. If the obtained system is inconsistent, the equations involving complemented bits of the intercepted sequence are eliminated from the system and then the consistency of the system is checked again. Which bits of the intercepted sequence are complemented is also guessed. If the guess of those bits is right, the obtained system will surely become consistent. Otherwise, if the guess of the initial state of the clocking LFSR was right, there is a significant chance that the newly obtained system becomes consistent. In a contrary case, the new system will be inconsistent with high probability. The attack always gives the right solution for the initial state of the clocking LFSR, but that solution may be accompanied by other solutions (false alarms). The recovery procedure in those cases is proposed as well. The attack is feasible if the number of possible combinations for the bits of the intercepted

sequence complemented by noise is small, which means that the level of noise is up to moderate.

REFERENCES

- [1] G. Bu, "Linear consistency test (LCT) in cryptanalysis of irregularly clocked LFSRs in the presence of noise," Master thesis, Gjøvik University College, Gjøvik, Norway, 2011.
- [2] W. Chambers and S. Jennings, "Linear equivalence of certain BRM shift-register sequences," *Electronics Letters*, vol. 20, no. 24, pp. 1018–1019, 1984.
- [3] J. Golić and M. Mihaljević, "A generalized correlation attack on a class of stream ciphers based on the Levenshtein distance," *Journal of Cryptology*, vol. 3, no. 3, pp. 201–212, 1991.
- [4] T. Johansson, "Reduced complexity correlation attacks on two clock-controlled generators," in: Ohta K. (Ed.), *Advances in Cryptology: Proceedings of ASIACRYPT '98*, Lecture Notes in Computer Science LNCS 1514, pp. 342–356, Springer-Verlag, 1998.
- [5] H. Molland, T. Helleseeth, "An improved correlation attack against irregular clocked and filtered keystream generators," in *Proceedings of CRYPTO 2004*, Lecture Notes in Computer Science LNCS 3152, pp. 373–389, Springer-Verlag, 2004.
- [6] S. Petrović, "Application of linear consistency test in a ciphertext-only attack on irregularly clocked linear feedback shift registers," in *Proceedings of XII Spanish Conference on Cryptography and Information Security (RECSI2012)*, U. Zurutuza, R. Uribeetxeberria, I. Arenaza-Nuño, Eds. Arrasate - Mondragon: Servicio Editorial de Mondragon Unibertsitatea, pp. 113–117, 2012.
- [7] K. Zeng, C. Yang, and T. Rao, "On the linear consistency test (LCT) in cryptanalysis with applications," in *Advances in Cryptology, Proceedings of CRYPTO '89*, Lecture Notes in Computer Science LNCS 435, pp. 164–174, Springer-Verlag, 1990.