

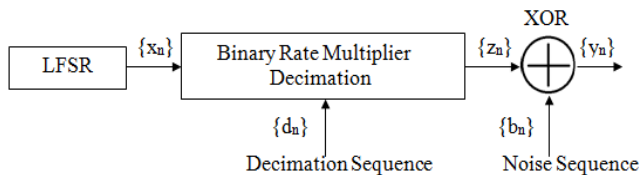


Given the prefix of the intercepted output sequence of length  $M$ , determine the initial states of all the LFSRs in the cascade. The goal of the paper is to determine how this task could be solved.

The decimation of sequences happens when the output sequence of a subgenerator is fed into the clock control input of one or more other subgenerators. The minimal example of such a scheme is a scheme in which one LFSR clocks another. In this paper, we generalize the ideas from Ref. [14] to cryptanalysis of cascades of irregularly clocked LFSRs.

## 2. A PARTICULAR STATISTICAL MODEL APPLIED IN CRYPTANALYSIS OF CASCADE OF LFSRS

The statistical model that we are interested in is based on the binary rate multiplier. This model was originally mentioned in Ref. [15]. The model is applied in Ref. [14] for the cryptanalysis of pseudorandom sequence generators. It is shown in Figure 2. In this paper, we analyze a stage of the cascade using this model.



**Figure 2.** The statistical model of a stage of the cascade

According to Golić and Mihajević's research [15], the feedback polynomial of the LFSR in this model is given as:

$$f(x) = 1 + a_1x + a_2x^2 + \dots + a_nx^L \quad (1)$$

where,  $L$  is the given length of this LFSR.  $\{x_n\}$  is the output sequence of the LFSR.  $\{d_n\}$ , serving as the decimation sequence, is the output sequence from another LFSR.  $\{z_n\}$  is the output sequence after the decimation processor.  $\{b_n\}$  is the binary noise sequence, for example, the plaintext.  $\{y_n\}$  is produced by the sum modulo 2 of the decimated sequence  $\{z_n\}$  and the noise sequence  $\{b_n\}$ .

$$y_n = z_n \oplus b_n \quad (2)$$

$$z_n = x_{f(n)} \quad (3)$$

$$f(n) = n + \sum_{i=0}^n d_i, n = 0, 1, 2, \dots \quad (4)$$

In this statistical model, it is supposed that  $\{d_n\}$  is the realization of the sequence  $\{D_n\}$  of independent and identically distributed random variables, with a certain probability  $Pr(D_n=i)$ , which should be chosen in our research as it is described in Section 3.

The binary noise sequence  $\{b_n\}$  is the realization of the sequence of random independent and identically distributed variables  $\{B_n\}$  with probability:

$$Pr(B_n = 1) = p < 0.5, \forall n \quad (5)$$

## 3. CRYPTANALYSIS OF THE CASCADE WITH AN ARBITRARY NUMBER OF LFSRS

### 3.1 The phases of cryptanalysis of cascade with two LFSRs

Two phases of cryptanalysis of cascade with two LFSRs (Figure 1) are discussed in Ref. [14]. In this cascade, there are two LFSRs. The first register LFSR<sub>1</sub> is regularly clocked while the second one LFSR<sub>2</sub> is irregularly clocked with the clock signals from the output sequence of LFSR<sub>1</sub> (Figure 1). The clocking satisfies the model described in Figure 2.

As it can be seen in Figure 2, the output sequence of LFSR<sub>1</sub>  $\{d_n\}$  serves as the decimation sequence. The output sequence of LFSR<sub>2</sub>  $\{x_n\}$  is decimated by  $\{d_n\}$ . The output sequence of the binary rate multiplier  $\{z_n\}$  and the noise sequence  $\{b_n\}$  are summed modulo 2 to produce the final output sequence  $\{y_n\}$ .

There are two main phases in the cryptanalysis of the cascade with two LFSRs. The two phases are briefly described in Ref. [14], in which the emphasis is put on the reconstruction of suboptimal paths in the constrained edit distance array.

The first phase of the cryptanalysis of a cascade with two LFSRs is that the candidate initial states of LFSR<sub>2</sub> should be determined. Golić and Mihajević [15] describes a correlation attack based on the statistical model given in Section 2. Firstly, certain number of initial states are chosen at random from all the non-zero initial states of LFSR<sub>2</sub>. Then, these selected initial states are used to generate  $\{x_n\}$ , whose length depends on the length of the intercepted sequence. After that, the constrained edit distance between the intercepted sequence and each generated  $\{x_n\}$  is calculated. The threshold is chosen to be an integer greater than or equal to the maximum of the results of computation of the constrained edit distance. In the next step, we use all the remaining initial states of LFSR<sub>2</sub> to generate  $\{x_n\}$ , between which and the intercepted sequence, a new set of constrained edit distances is calculated. The states producing  $\{x_n\}$ , whose constrained edit distance to the intercepted output sequence is less than the threshold represent the candidate initial states for LFSR<sub>2</sub>.

The second phase of the attack is to reconstruct the clock control sequence of LFSR<sub>2</sub>. In 2007, Petrovic and Fuster-Sabater [14] proposed an attack, in which a new algorithm was used to search for the optimal and suboptimal paths through the matrix of constrained edit distances, each corresponding to a possible clock control sequence for LFSR<sub>2</sub>. This search is directed by increasing the tolerance for weight discrepancy between the weight (i.e. constrained edit distance) of a reconstructed suboptimal path and the weight of the optimal path. In this paper, we extend the ideas from Ref. [14] to cryptanalysis of a cascade of irregularly clocked LFSRs containing an arbitrary number of LFSRs.

### 3.2 The method of splitting the cascade into subcascades

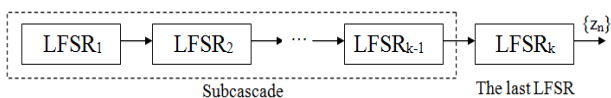
The scheme of a cascade of LFSRs presented can be divided into two parts: the first one is the subcascade that generates the clock control sequence for the last LFSR, and the second one is the last LFSR in the cascade itself (as shown in Figure 3 below).

If the prefix of the output sequence of the last LFSR is known to the attacker, it is possible to reconstruct the initial state of the last LFSR with a generalized correlation attack. It is possible to obtain a set of candidate initial states of that LFSR, which could generate the intercepted output sequence. This can be realized by using the model described in Section

2.

The next step is to determine which initial state, among all the candidate initial states obtained, can generate the intercepted output sequence. So the control sequence that, together with one of the candidate initial states of the last LFSRs in the cascade, could generate the prefix of the intercepted sequence has to be determined. We use the constrained edit distance matrix associated with each candidate initial state to reconstruct the paths, each of which maps to a possible control sequence.

After the possible clock control sequences of the last LFSR in the cascade are reconstructed, we can treat the rest of the scheme as a cascade of lower dimension. We can use the possible clock control sequences of the last LFSR as the output sequence of the subcascade. Then the same process can be used in the subcascade.



**Figure 3.** The scheme of a cascade of irregularly clocked LFSRs divided into two parts

## 4. EXPERIMENTAL WORK

### 4.1 Objective of the experiment

The purpose of our experiment was to verify that the methods discussed in the previous sections can be used in cryptanalysis of a cascade of LFSRs. A similar experiment with two LFSRs was implemented in Ref. [14]. In our experiment, by using the theory discussed in the previous sections, we carry out cryptanalysis of a cascade with three LFSRs. By doing this, the method is extended from the cascade with two LFSRs to that with three LFSRs. Then the same procedure can be applied to extending to the cascade with an arbitrary number of LFSRs in the future work.

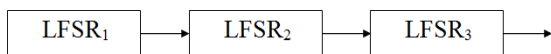
We chose three LFSRs constituting the cascade (see Figure 4). These LFSRs are clocked in the following way:

LFSR<sub>1</sub> is regularly clocked.

If the output of LFSR<sub>1</sub> is 0, LFSR<sub>2</sub> is clocked once, and the output bit of LFSR<sub>2</sub> is valid and is used to clock LFSR<sub>3</sub>.

If the output of LFSR<sub>1</sub> is 1, LFSR<sub>2</sub> is clocked twice. The first output bit of LFSR<sub>2</sub> is discarded. The second output bit is valid and is sent to clock LFSR<sub>3</sub>.

The same procedure is applied in clocking LFSR<sub>3</sub> by using the valid output bit of LFSR<sub>2</sub>.



**Figure 4.** A cascade of three LFSRs

In our experiment, as a matter of convenience, we chose three LFSRs with the same feedback polynomials:

$$f(x) = 1 + x^2 + x^7 + x^6 + x^{10} \quad (6)$$

Each LFSR had 4 feedback taps and 1 output tap. The output tap was the content of the leftmost cell of each LFSR. In addition, all the LFSRs had the length of 10.

## 4.2 Simulation of generating the intercepted sequence

The aim of the attack was, given the intercepted sequence (the output sequence of LFSR<sub>3</sub> with a certain level of noise), to determine the initial states of LFSR<sub>3</sub>, LFSR<sub>2</sub> and LFSR<sub>1</sub> successively as well as the correct clocking sequences of LFSR<sub>3</sub> and LFSR<sub>2</sub>. In order to demonstrate our attack procedure to be feasible, we needed to compare the initial states and the clocking sequences we obtained from our attack to the correct ones, which were known in advance. This could be realized by the following logical steps:

1. The initial states of each LFSR in the cascade were chosen at random. They were recorded.
2. With the initial states obtained in the step 1 and the clocking algorithm, we could start the simulation procedure. By simulating the irregular clocking procedure, the intercepted sequence was generated.
3. The corresponding output sequences of LFSR<sub>2</sub> and LFSR<sub>1</sub>, which were used to generate the intercepted sequence, were recorded.
4. The attack (cryptanalysis) procedure was started bearing in mind the fact that the only known information was the intercepted sequence and the feedback polynomials of each LFSR.
5. Compare the results of the attack, including the initial states of the three LFSRs and the output sequences of LFSR<sub>3</sub> and LFSR<sub>2</sub>, to the recorded ones in step 1 and 3.
6. If the values of each comparison are identical, the attack method is proved to be feasible.

In the simulation procedure of our experiment, the initial states of LFSR<sub>1</sub>, LFSR<sub>2</sub> and LFSR<sub>3</sub> were set to 0101111111, 0011100100 and 1011100000, respectively. The length of the output intercepted sequence was set to 200. The LFSRs had feedback polynomials (1). Then the simulation was executed and we generated 200 bits of the intercepted sequence Y as well as the corresponding output sequence of LFSR<sub>1</sub> (clock 1) and LFSR<sub>2</sub> (clock 2), and the intercepted sequence (Y<sub>1</sub>) without noise.

### 4.3 Cryptanalysis processes and results

The cryptanalysis (attack) procedure consists of the following steps:

1. Determine the candidate initial states of LFSR<sub>3</sub>.
2. Determine the correct clocking sequence of LFSR<sub>3</sub>.
3. Determine the candidate initial states of LFSR<sub>2</sub>.
4. Determine the correct clocking sequence of LFSR<sub>2</sub>.
5. Determine the correct initial state of LFSR<sub>1</sub>.

In this cryptanalysis procedure, it is supposed that the feedback polynomials are known in advance as shown in polynomials (1). In the first place, the method used in the realization of step 1 and step 3 was described in Ref. [15]. Furthermore, after we obtained the clocking sequence of LFSR<sub>2</sub> in the step 4, we can calculate the initial state of LFSR<sub>1</sub> by solving the system of linear equations in the step 5. So, in our experiment, we shall only focus on the step 2 and the step 4.

In the step 1, with the knowledge of the intercepted sequence, we can determine a set of candidate initial states of LFSR<sub>3</sub> (see the correlation attack described in Ref. [15]). In our experiment, we supposed that we obtained the correct initial state of LFSR<sub>3</sub> which, together with the corresponding clocking sequence of LFSR<sub>3</sub>, could produce the intercepted sequence.

In the step 2, the initial state of LFSR<sub>3</sub> was set to 1011100000, which was equal to the value used in the simulation process. And the intercepted sequence used to establish the constrained edit distance was also set to be equal to the value obtained by the simulation process. The length of the intercepted sequence was set to 200. Besides, the length pl out of the clocking sequence of LFSR<sub>3</sub> was set to 40, which means only the first 40 bits of clock 2 were used in the comparison process. Then we generated the corresponding constrained edit distance matrix and reconstructed all the optimal and sub-optimal paths using algorithm in Ref. [16]. The result shows that, after searching 7612 paths, the correct clock sequence is found equal to clock 2.

The same process was applied in the step 4, where the initial state of LFSR<sub>2</sub> was set to 0011100100. As the feedback polynomial of LFSR<sub>1</sub> was known, the length pl was set to 10, which was equal to the length of LFSR<sub>1</sub>. The length of clock 2 needed was set to 40. Then we generated the corresponding constrained edit distance matrix and reconstructed the optimal and sub-optimal paths using algorithm in Ref. [16] again. The result shows that, after searching only 20 paths, the correct clock-control sequence was found equal to clock 1.

The experiment results proved that the theory discussed in previous sections and the methods described in this section were feasible.

## 5. CONCLUSION

In this paper, the cryptanalysis of the cascade of irregularly clocked linear feedback shift registers is described. The attack is essentially a method for reconstruction of the initial state of the subcascade that generates the clock control sequence. The statistical model based on binary rate multiplier, which employs the constrained edit distance is used. Instead of checking all the possible initial states of the subcascade, all the possible optimal paths in the edit distance matrix as well as the suboptimal paths, whose weight-difference from the optimal ones dose not overcome the discrepancy D given in advance, are reconstructed by depth-first search. Experimental results show that the methods described in this paper, which is used to cryptanalyze a cascade with an arbitrary number of LFSRs, are applicable.

The experiment results proved that it is feasible to generalize the correlation attack against a scheme with 2 LFSRs, of which one irregularly clocks another, to a cascade of irregularly clocked LFSRs. First of all, the cascade with an arbitrary number of LFSRs was split into subcascades. Then, the possible candidate initial states of the last LFSR were reconstructed with a generalized correlation attack. Once the set of candidate initial states is known, the element that generates the intercepted output sequence has to be determined. The attack continued by determining the clock control sequence that, together with one of the candidate initial states of the last LFSR, could generate the intercepted sequence. This can be achieved by searching through the constrained edit distance matrix associated with every candidate initial state obtained in the previous step of attack. The optimal and suboptimal paths were reconstructed in the searching procedure. Once the possible clock control sequences of the last LFSR in the cascade were reconstructed, we treated the rest of the scheme as a cascade of lower dimension utilizing the reconstructed clock sequences as its output sequence. Then, the same analytical methods were applied into the cascade of

lower dimension. This process continues until the possible initial states of the first LFSR in the cascade were reconstructed.

## REFERENCES

- [1] Beth, T., Piper, F. (1984). The stop-and-go-generator. Proceeding of EUROCRYPT 84, LNCS 209, Springer Verlag, Berlin, 88-92. [https://doi.org/10.1007/3-540-39757-4\\_9](https://doi.org/10.1007/3-540-39757-4_9)
- [2] Chambers, W.G., Jennings, S.M. (1984). Linear equivalence of certain BRM shift-register sequences. Electronics Letters, 10(24): 1018-1019. <https://doi.org/10.1049/el:19840693>
- [3] Coppersmith, D., Krawczyk, H., Mansour, Y. (1993). The shrinking generator. Advances in Cryptology — CRYPTO' 93, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 773: 22-39. [https://doi.org/10.1007/3-540-48329-2\\_3](https://doi.org/10.1007/3-540-48329-2_3)
- [4] Günther, C.G. (1987). Alternating step generators controlled by de Bruijn sequences. EUROCRYPT'87 Proceedings of the 6th annual international conference on Theory and application of cryptographic techniques. Springer Verlag, Berlin, Heidelberg, pp. 5-14. [https://doi.org/10.1007/3-540-39118-5\\_2](https://doi.org/10.1007/3-540-39118-5_2)
- [5] Menezes, A., VanOorschot, P., Vanstone, S. (1997). Handbook of applied cryptography. CRC Press. <http://dx.doi.org/10.1201/9781439821916>
- [6] Rueppel, R.A. (1986). Analysis and design of stream ciphers. Springer-Verlag. <https://doi.org/10.1007/978-3-642-82865-2>
- [7] Beker, H., Piper, F. (1982). Cipher systems: The protection of communications. Northwood Books. <https://doi.org/10.1112/blms/15.5.521>
- [8] Krause, M., Armknecht, F. (2003). Algebraic attacks on combiners with memory. Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA. [https://doi.org/10.1007/978-3-540-45146-4\\_10](https://doi.org/10.1007/978-3-540-45146-4_10)
- [9] Courtois, N. (2003). Fast algebraic attacks on stream ciphers with linear feedback. Proceedings of Crypto 2003, LNCS 2729, Springer Verlag, pp. 176-194. [https://doi.org/10.1007/978-3-540-45146-4\\_11](https://doi.org/10.1007/978-3-540-45146-4_11)
- [10] Siegenthaler, T. (1985). Decrypting a class of stream ciphers using ciphertext only. IEEE Transactions on Computers, C34(1): 81-85. <http://dx.doi.org/10.1109/TC.1985.1676518>
- [11] Gollman, D., Chambers, W.G. (1989). Clock-controlled shift registers: A review. IEEE Journal on Selected Areas in Communications, 7(4): 525-533. <https://doi.org/10.1109/49.17716>
- [12] Oommen, B.J. (1986). Constrained string editing. Elsevier Science Inc, 40(3): 267-284. [https://doi.org/10.1016/0020-0255\(86\)90061-7](https://doi.org/10.1016/0020-0255(86)90061-7)
- [13] Oommen, B.J. (1987). Recognition of noisy subsequences using constrained edit distance. IEEE Transactions on Pattern Analysis & Machine Intelligence, 9(5): 676-685. <https://doi.org/10.1109/TPAMI.1987.4767962>
- [14] Petrović, S., Fúster, A. (2007). Reconstruction of suboptimal paths in the constrained edit distance array with application in cryptanalysis with application in cryptanalysis. Proceedings of the 2007 international

- conference on computational science and its applications, ICCSA 2007, Kuala Lumpur, Malaysia, Lecture Notes in Computer Science, LNCS 4707, Part III: 597-610. [http://dx.doi.org/10.1007/978-3-540-74484-9\\_52](http://dx.doi.org/10.1007/978-3-540-74484-9_52)
- [15] Golić, J., Mihaljević, M. (1991). A generalized correlation attack on a class of stream ciphers based on the Levenshtein distance. *Journal of Cryptology*, 3(3): 201-212. <https://doi.org/10.1007/BF00196912>
- [16] Petrović, S., Fúster, A. (2004). Clock control sequence reconstruction in the ciphertext only attack scenario. *International Conference on Information and Communications Security, Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, 3269: 427-439. [https://doi.org/10.1007/978-3-540-30191-2\\_33](https://doi.org/10.1007/978-3-540-30191-2_33)