A COMPARISON OF CRYPTANALYTIC PRINCIPLES BASED ON ITERATIVE ERROR-CORRECTION

Miodrag J. Mihaljević and Jovan Dj. Golić

Institute of Applied Mathematics and Electronics, Belgrade
School of Electrical Engineering, University of Belgrade
Bulevar Revolucije 73, 11001 Beograd, Yugoslavia

ABSTRACT: A cryptanalytic problem of a linear feedback shift register initial state reconstruction using a noisy output sequence is considered. The main underlying principles of three recently proposed cryptanalytic procedures based on the iterative error-correction are pointed out and compared.

I. INTRODUCTION

A weakness of a class of running key generators for stream ciphers is demonstrated in [1], and fast algorithms for the cryptanalysis are proposed in [2]-[7] having origins in [8]. In this paper the main underlying principles for the algorithms [2]-[6] are analyzed. The following three principles are considered:

P.1: Error-correction is based on the number of satisfied parity-checks.

P.2: Error-correction is based on the estimation of the relevant posterior probabilities obtained by using the average posterior probability estimated in the previous iteration as the prior probability in the current iteration.

P.3: Error-correction is based on the estimation of the relevant posterior probabilities obtained by using the posterior probabilities estimated in the previous iteration as the prior probabilities in the current iteration.

II. ALGORITHMS

In this section three algorithms corresponding to the principles P.1-P.3 are specified. Algorithm P.1 is the algorithm proposed in [3]. Algorithm P.2 could be regarded as a simplification of the Algorithm [4]. Algorithm P.3 could be seen as a simplification/modification of the Algorithm B [2].

Denote by $\{x_n\}_{n=1}^N$ an output segment of a linear feedback shift register (LFSR) of length L with w feedback tapes. In a statistical model, a binary noise sequence $\{e_n\}_{n=1}^N$ is assumed to be a

D.W. Davies (Ed.): Advances in Cryptology - EUROCRYPT '91, LNCS 547, pp. 527-531, 1991.

realization of a sequence of i.i.d. binary variables $\{E_n\}_{n=1}^N$ such that $\Pr(E_n=1) = p$, $n=1,2,\ldots,N$. Let $\{z_n\}_{n=1}^N$ be a noisy version of $\{x_n\}_{n=1}^N$ defined by

$$z_n = x_n \oplus e_n \quad , \quad n=1,2,\ldots,N \quad . \tag{1}$$

The problem under consideration is a reconstruction of the LFSR initial state based on the principles P.1-P.3 assuming that the segment $\{z_n\}_{n=1}^N$, the LFSR characteristic polynomial, and the parameter p are known. For the comparison purposes we assume that all the algorithms are based on the parity-checks defined as follows.

Definition: $\Pi_n = \{\pi_k(n)\}_k$ is a set of orthogonal parity-checks related to the n-th bit that are generated according to the characteristic polynomial multiples as in [2]-[3], $n=1,2,\ldots,N$.

Let

$$c_k(n) = \sum_{\ell\in\pi_k(n)} {}_{\mathrm{mod}2}\, z_\ell \quad , \quad k=1,2,\ldots,|\Pi_n| \quad , \quad n=1,2,\ldots,N \quad . \tag{2}$$

where $|\Pi_n|$ denotes the cardinality of Π_n. Assume that $c_k(n)$ is a realization of a binary random variable $C_k(n)$, $k=1,2,\ldots,|\Pi_n|$, $n=1,2,\ldots,N$. Let $\Pr(E_n, \{C_k(n)\}_{k=1}^{|\Pi_n|})$ be the joint probability of the variables E_n and $C_k(n)$, $k=1,2,\ldots,|\Pi_n|$, and let $\Pr(E_n | \{C_k(n)\}_{k=1}^{|\Pi_n|})$ be the corresponding posterior probability, $n=1,2,\ldots,N$.

The following steps are identical for all the algorithms:

Initialization: $i=0$, $I=\mathrm{const}$, $p^{(0)}=p$.

Step 1: Set $i \rightarrow i+1$. If $i > I$ go to the last step.

Step 2: Calculate $c_k(n)$, $k=1,2,\ldots,|\Pi_n|$, $n=1,2,\ldots,N$.

ALGORITHM P.1 [3]:

Step 3: Calculate $t_n = |\Pi_n| - 2\sum_{k=1}^{|\Pi_n|} c_k(n)$, $n=1,2,\ldots,N$.

Step 4: If $t_n < 0$, set $z_n \rightarrow z_n \oplus 1$, $n=1,2,\ldots,N$. Go to Step 1.

Step 5: Stop the procedure.

ALGORITHM P.2:

Step 3: For $n=1,2,\ldots,N$, calculate

$$p_n^{(i)} = \Pr(E_n=1 \,|\, \{C_k(n)\}_{k=1}^{|\Pi_n|} = \{c_k(n)\}_{k=1}^{|\Pi_n|}) =$$

$$\frac{p^{(i)}\, p_w^{s_n}\, (1-p_w)^{|\Pi_n|-s_n}}{p^{(i)}\, p_w^{s_n}\, (1-p_w)^{|\Pi_n|-s_n} + (1-p^{(i)})\, (1-p_w)^{s_n}\, p_w^{|\Pi_n|-s_n}} \quad . \tag{3}$$

where

$$s_n = \Sigma_{k=1}^{|\Pi_n|} c_k(n) \quad , \quad p_w = [1-(1-2p^{(i)})^w] \,/\, 2 \quad . \tag{4}$$

Step 4: If $P_n^{(i)} > 0.5$, set $z_n \to z_n \oplus 1$, $P_n^{(i)} \to 1-P_n^{(i)}$, $n=1,2,\ldots,N$.

Step 5: Calculate $p^{(i)} = (1/N) \sum_{n=1}^{N} P_n^{(i)}$. Go to Step 1.

Step 6: Stop the procedure.

ALGORITHM P.3:

Step 3: Calculate

$$P_n^{(i)} = \Pr(E_n=1 \,|\, \{C_k(n)\}_{k=1}^{|\Pi_n|} = \{c_k(n)\}_{k=1}^{|\Pi_n|}) =$$

$$\frac{p_n^{(i)} \prod_{\ell=1}^{|\Pi_n|} p_\ell(n)^{c_\ell(n)} [1-p_\ell(n)]^{\bar{c}_\ell(n)}}{p_n^{(i)} \prod_{\ell=1}^{|\Pi_n|} p_\ell(n)^{c_\ell(n)} [1-p_\ell(n)]^{\bar{c}_\ell(n)} + (1-p_n^{(i)}) \prod_{\ell=1}^{|\Pi_n|} [1-p_\ell(n)]^{c_\ell(n)}\, p_\ell(n)^{\bar{c}_\ell(n)}} \tag{5}$$

where

$$\bar{c}_\ell(n) = 1 - c_\ell(n) \quad , \quad p_\ell(n) = [1 - \prod_{j=1}^{w} (1 - 2\, p_{m_j})] \,/\, 2 \quad . \tag{6}$$

and $\{m_j\}_{j=1}^{w}$ denotes the set of indices of the bits involved in the parity-check $\pi_\ell(n)$, for any $\ell=1,2,\ldots,|\Pi_n|$, $n=1,2,\ldots,N$.

Step 4: If $P_n^{(i)} > 0.5$, set $z_n \to z_n \oplus 1$, $P_n^{(i)} \to 1-P_n^{(i)}$, $n=1,2,\ldots,N$.

Step 5: Set $p_n^{(i)} \to P_n^{(i)}$, $n=1,2,\ldots,N$. Go to Step 1.

Step 6: Stop the procedure.

III. EXPERIMENTAL RESULTS

The experiments are realized using an LFSR of length 47 with 2 feedback tapes on the stages 5 and 47, when the observed sequence is of length $N=10^5$. The following self-explanatory table presents the experimental results. According to the experimental investigations, all the algorithms could work when the noise is under a limit which is a function of the observed sequence length. For higher noise, Algorithm P.1 is the first to fail, and Algorithm P3 is the last one to fail.

Table: The number of residual errors as a function of the iteration step for Algorithms P.1-P.3 and the noise $p = p_1 \cdot p_2 \cdot p_3$ where $p_1=0.400$, $p_2=0.425$ and $p_3=0.435$.

iteration	# of residual errors								
i	Algorithm P.1			Algorithm P.2			Algorithm P.3		
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3
1	40357	44440	45774	37728	41693	43077	37728	41693	43077
2	40383	45868	47301	35734	41397	43015	34462	40943	42712
3	39343	46758	48388	33477	41002	42934	30249	40194	42397
4	36610	47147	48566	30400	40659	42814	24943	39270	42211
5	31750	47468	48763	26130	40259	42821	15333	38191	41977
6	23614	47779	48626	19808	39827	42657	5719	36618	41796
7	13714	47610	48699	11850	39214	42522	1484	34849	41376
8	6246	47530	48817	6315	38544	42423	117	32711	41133
9	1820	47736	48667	3184	38935	42359	2	30097	40768
10	230	47606	48699	717	38661	42335	0	26603	40515
11	0	47528	48704	13	38432	42347		22190	40156
12		47574	48820	0	38216	42346		16766	39918
13		47478	48962		38028	42326		11810	39579
14		47532	48854		37870	42337		8403	39307
15		47551	48878		37688	42315		6110	39033
16		47466	48822		37505	42344		4006	38755
17		47578	48852		37320	42344		2198	38420
18		47613	48623		37127	42358		831	38079
19		.	48790		36940	42348		139	37718
20		.	48704		36661	42340		0	37277
21		.	48800		36304	42338			36800
22		.	48776		35838	42340			36235
23		.	48785		35225	42343			35655
24		.	48763		34429	42349			35003
25		.	48862		33569	42351			34262
26		.	48762		32504	42356			32350
27		.	48835		31189	42350			31183
28		.	48818		29703	42353			29750
29		.	48893		28146	42355			28273
30		.	48805		26409	42352			25309
31		.	48833		24191	42352			23818
32		.	48816		21280	42352			22280
33		.	48835		18105	42358			20518
34		.	48789		15042	42360			18441
35		.	48801		12245	42360			15922
36		.	.		9443	42360			12801
37		.	.		7080	42360			9685
38		.	.		5197	42360			7140
39		.	.		3446	42360			5337
40		.	.		1910	42360			3837
41		.	.		745	42360			2604
42		.	.		122	42360			1317
43		.	.		0	.			329
44		.	.			.			3
45		.	.			.			0

IV. CONCLUSIONS

A cryptanalytic problem of an LFSR initial state reconstruction using the noisy output sequence is considered. The main underlying principles of the cryptanalytic algorithms based on the iterative error-correction, recently proposed in [2]-[6], are compared. The three corresponding algorithms, named Algorithms P.1-P.3, are specified and analyzed.

Let an iteration cost be an equivalent of the iteration cycle complexity and a reconstruction cost be a product of the iteration cost and the number of iterations needed for the reconstruction. The main complexity difference between the algorithms is in the third step. Note that, for a given $|\Pi_n|$, the probability (3) depends only on $s_n=\Sigma_{k=1}^{|\Pi_n|} c_k(n)$, instead of the individual parity-checks $c_k(n)$. Accordingly, it can be shown that the complexity of Algorithm P.3 is considerably greater than the complexities of both Algorithms P.1 or P.2.

According to the experimental results and the complexity analysis, we have the following heuristic conclusions:

- When the noise is lower than the limit below which all the algorithms work, Algorithm P.1 yields the minimum reconstruction cost.
- In the case of higher noise when Algorithm P.1 fails and both Algorithms P.2 and P.3 work, it is better to use Algorithm P.2 because of the lower reconstruction cost.
- Finally, when Algorithm P.3 works and Algorithms P.1 and P.2 both fail, in order to minimize the reconstruction cost the following procedure could be used: make the initial error-rate reduction using Algorithm P.3, and after the certain points change the running algorithm by Algorithms P.2 and P.1, respectively.

REFERENCES

[1] T.Siegenthaler, "Decrypting a Class of Stream Ciphers Using Ciphertext Only", IEEE Trans. Comput., vol. C-34, Jan. 1985, pp.81-85.

[2] W.Meier, O.Staffelbach, "Fast Correlation Attacks on Certain Stream Ciphers", Journal of Cryptology, vol.1, 1989., pp.159-176.

[3] K.Zeng, M.Huang, "On the Linear Syndrome Method in Cryptanalysis", Lecture Notes in Computer Science, Advances in Cryptology - CRYPTO '88, vol.405, pp.469-478. Springer-Verlag, 1990.

[4] M.Mihaljević, J.Golić, "A Fast Iterative Algorthm for a Shift Register Initial State Reconstruction Given the Noisy Output Sequence", Lecture Notes in Computer Science, Advances in Cryptology - AUSCRYPT '90, vol.453, pp.165-175. Springer-Verlag, 1990.

[5] K.Zeng, C.H.Yang, T.R.N.Rao, "An Improved Linear Syndrome Algorithm in Cryptanalysis with Applications", to appear in Lecture Notes in Computer Science, Advances in Cryptology - CRYPTO '90.

[6] V.Chepyzhov, B.Smeets, "On a Fast Correlation Attack on Stream Ciphers", EUROCRYPT "91.

[7] M.Živković, "An Analysis of Linear Recurrent Sequences over the Field GF(2)", Ph.D. thesis, University of Belgrade, 1990.

[8] R.G.Gallager, "Low-Density Parity-Check Codes", IRE Trans. Inform. Theory, vol. IT-8, Jan. 1962, pp.21-28.