Domarev Dmitry, postgraduate student of the National aviation university by speciality 21.05.01 «Information security of the State».

Милокум Яна Валеріївна, к.т.н., доцент, Національний авіаційний університет,
E-mail: Yana.Mylokum@gmail.com
Милокум Яна Валериевна, к.т.н., доцент, Национальный авиационный университет.
Mylokum Yana, PhD in technical sciences, Assossiate Professor, National aviation university.

UDC 621.391:519.7

ON THE COMPUTATIONAL SECURITY OF RANDOMIZED STREAM CIPHERS PROPOSED BY MIHALJEVIĆ AND IMAI

Anton Alekseychuk, Sergey Gryshakov

This paper yields a (computational) security analysis for a generic class of randomized stream ciphers based on joint employment of encryption, error-correction coding, and dedicated random coding. We show that the security of these ciphers can be considerably less than their designers claim. In contrast to the approach for security evaluation used before, our technique is significantly simpler and allows us to find out the code-theoretic sense of parameters that determine the security of these ciphers. We also propose another possible solution (based on nonlinear random coding) for design of randomized stream ciphers with enhanced security.
Keywords: *symmetric cryptography, randomized encryption, stream cipher, random coding, wiretap channel, LPN problem, correlation attack.*

1. Introduction

M. J. Mihaljević and H. Imai [8, 9, 11, 12] proposed a general approach for design of randomized stream ciphers based on joint employment of encryption, error-correction coding, and dedicated random (or homophonic) coding. One of the goals of designing such ciphers is to increase the security (without substantial performance reducing) of stream ciphers currently used in wireless communication systems, particularly, in the GSM standard. Another reason is to construct symmetric encryption schemes, whose security can be reduced to the hardness of some known mathematical problem such as the *Learning from Parity with Noise* (LPN) problem. Recall (see [6], for example) that this problem consists in solving a system of linear Boolean equations with equiprobable random coefficient matrix and the right-hand side corrupted by independent random variables taking values 0 and 1 with probabilities $1-\theta$ and θ, respectively, $\theta \in (0, 1/2)$. In this case, we say that θ is the *noise level* in the right-hand side of the given system of linear equations.

In what follows, we focus our attention on the versions of randomized stream ciphers defined in [11, 12] and studied in detail in [10, 11, 13].

Let's denote by V_n the set of all n-dimensional Boolean vectors, by $F_{m\times n}$ the set of $m \times n$-matrices over the field $F = \mathbf{GF}(2)$, and by $F^*_{m\times m}$ the group of all invertible matrices of order m over this field.

According to [11, 12], the initial objects for a randomized stream cipher with parameters $l, m, n \in \mathbf{N}$, $p \in (0, 1/2)$, where $l < m < n$, and a key space K are matrices $G_1 \in F_{m\times n}$, $G_2 \in F^*_{m\times m}$, and a keystream generator that produces a sequence $f_0(k), f_1(k), \ldots$ of n-dimensional Boolean vectors determined by a key $k \in K$. It is assumed that the functions $f_i : K \to V_n$, $i = 0, 1, \ldots$, can depend on some public parameters, for example, on initialization vectors (IV's). It is also assumed that G_1 is a generator matrix of a binary linear $[n, m]$-code C_1 with an efficient decoding algorithm, which is guaranteed to correct errors in the binary symmetric channel with crossover probability p.

To encrypt a plaintext $s_0, s_1, \ldots, s_t$, where $s_i \in V_l$, $i = 0, 1, \ldots, t$, with a key $k \in K$ the sender generates a sequence of independent random vectors $u_0, v_0, u_1, v_1, \ldots, u_t, v_t$, where u_i is uniformly distributed on the set V_{m-l}, and v_i is distributed according to Bernoulli's law with parameters (n, p), and computes the ciphertext $z_0, z_1, \ldots, z_t$ as follows:

$$z_i = (s_i, u_i)G_2G_1 \oplus f_i(k) \oplus v_i, \; i = 0, 1, \ldots, t. \quad (1)$$

The legitimate receiver, knowing $f_i(k)$, can quickly find the message $(s_i, u_i)G_2$ with the efficient decoding algorithm for the code C_1, after that he can recover s_i using invertibility of the matrix G_2. On the other hand, the adversary in order to find the key k will be forced to deal with a corrupted keystream $f_i(k) \oplus (s_i, u_i)G_2G_1 \oplus v_i$, $i = 0, 1, ..., t$.

In [10, 11, 13] different variants of specified randomized ciphers are investigated, in particular, with the following functions f_i:

$$f_i(k) = kS^i,\ k \in K = V_n, \tag{2}$$

where S is a non-secret $n \times n$-matrix over the field F;

$$f_i(k) = \alpha_i k,\ k \in K = F_{n \times n}, \tag{3}$$

where $\alpha_0, \alpha_1, ...$ are independent equiprobable random Boolean vectors of size n. Note that in the last case, a ciphertext is by definition the sequence $(\alpha_0, z_0), (\alpha_1, z_1), ..., (\alpha_t, z_t)$, where z_i are computed by formula (1). Strictly speaking, randomized encryption schemes of this type do not belong to the class of stream ciphers, and are proposed in [11] in order to generalize and to enhance one of the earlier probabilistic private-key encryption schemes, whose security can be reduced to the hardness of the LPN problem [6].

Based on the condition of implementation simplicity of the described encryption scheme, it was proposed in [10] to set up the matrices G_1 and G_2 as follows:

$$G_1 = \begin{pmatrix} I_{m-l} & 0 & A_1 \\ 0 & I_l & A_2 \end{pmatrix},\ G_2 = \begin{pmatrix} 0 & I_l \\ I_{m-l} & B \end{pmatrix} \tag{4}$$

where $A_1 \in F_{(m-l)\times(n-m)}$, $A_2 \in F_{l\times(n-m)}$, $B \in F_{(m-l)\times l}$, and I_l, I_{m-l} are identity matrices of specified size. In this case the transform

$$s \mapsto (s, u)G_2G_1 = (u, s \oplus uB, sA_2 \oplus u(A_1 \oplus BA_2)),$$
$$s \in V_l,\ u \in V_{m-l}, \tag{5}$$

used in (1) describes well-known combined random coding scheme for the wiretap channel proposed in fact in the fundamental paper [15] and extensively studied later (see [1, 14] for a comprehensive survey). Security evaluation of randomized ciphers specified by (1), both from information-theoretic and computational points of view, was performed in [10, 11, 13]. In [11], the authors claim that under condition (3) the secret key recovery of the proposed encryption scheme in the chosen plaintext attacking (CPA) scenario is as hard as solving the LPN problem with the noise level $1/2 \cdot \left(1-(1-2p)^{(m-l)/2}\right)$. Note that in the proof of this result an ad hock assumption about the matrix G_2G_1 is used; but, this assumption is not mentioned in the main part of the paper (see Proof of Theorem 2 in [11]).

In [10], the authors show that under conditions (1), (2), (4) the recovery of secret key in CPA scenario is reduced to the solving a system of linear Boolean equations corrupted by noise, where the noise level is at least $p_w = 1/2 \cdot \left(1-(1-2p)^{w+1}\right)$ and w depends somehow on the matrices G_1, G_2. In the same time, let us remark that condition $\mathrm{rank}(G^*) \geq w+1$ given in [10] (together with other conditions; see [10], p. 11) doesn't guarantee that the noise level in the right-hand side of obtained linear equations is lower bounded by p_w (this can be stated directly by analyzing an example given in [10], p. 15). Thus, the question about real computational security of ciphers proposed in [11, 12] is in fact open and requires further research.

Contribution of this Paper. We investigate the security of arbitrary ciphers specified by (1) and (4) in various attacking scenarios and show that it can be considerably less than it is claimed in [10 - 12]. In particular, we show that under condition (2) the complexity of the secret key recovery of the considered stream cipher in CPA scenario is upper bounded by the complexity of solving a system of linear Boolean equations corrupted by noise with the noise level p_w, where $w+1$ coincide with the dual distance of a code, determined by the matrices G_1 and G_2. Note that the specified system of equations has very dedicated form and can be solved considerably faster than the LPN problem with the same parameters. (Let us emphasize that this is not holds for the ciphers specified by (1) and (3). However, also in this case the noise level in the right-hand side of the obtained system of linear equations can be considerably less than the value $1/2 \cdot \left(1-(1-2p)^{(m-l)/2}\right)$ reported in [11]).

In contrast to the approach for security evaluation used in [10, 11], our technique is significantly simpler and allows us to find out the code-theoretic sense of parameters that determine the security of the considered ciphers. It follows from our results that to construct reasonably secure randomized encryption schemes, the matrices G_1 and G_2 should be very carefully selected, what seems a non-trivial problem. In concluding part of the paper we propose another possible solution (based on nonlinear random coding) for design of randomized stream cipher with enhanced security.

2. Preliminaries

In what follows, standard concepts and terminology from coding theory are used (see [4, 7] for more details). For a linear code $C \subseteq V_n$ the *dual code*, $C^{\perp}$, and the *dual distance*, $d(C^{\perp})$, of C are defined as follows:

$$C^{\perp} = \{y \in V_n \mid \forall x \in C : xy^T = 0\},$$

$$d(C^{\perp}) = \min\{wt(x) : x \in C^{\perp} \setminus \{0\}\},$$

where $wt(x)$ is the Hamming weight of a vector $x \in V_n$.

Let's consider an arbitrary cipher specified by (1), where matrices G_1 and G_2 are defined by (4). As above, let's denote by C_1 the linear code generated by the rows of the matrix G_1.

Set $C_0 = \{(0,u)G_2G_1 : u \in V_{m-l}\}$. It is clear that C_0 is an $[n, m-l]$-sub-code of the code C_1. We denote by $d_0^{\perp}$ and $d_1^{\perp}$ the dual distances of the codes C_0 and C_1, respectively. Note that the relation $C_0 \subseteq C_1$ implies the inequality

$$d_0^{\perp} \le d_1^{\perp}. \qquad (6)$$

Finally, for any $w = 0, 1, \ldots$ let's denote $p_w = 1/2 \cdot \left(1 - (1-2p)^{w+1}\right)$.

3. Main Results

Ciphertext-Only Attacks. Let us consider the system of equations (1) and suppose that adversary can access only the ciphertext $z_0, z_1, \ldots, z_t$. First of all, observe that, in order to recover plaintext symbols, the adversary is not required to have full information about the secret key.

Statement 1. *Let H be an arbitrary parity-check matrix of the code C_0; then for any $i = 0, 1, \ldots$ and $k \in K$ the adversary can recover (in real time) the vector s_i from the known values z_i and $\varphi_i(k) = f_i(k)H^T$.*

Proof. Let $a \in V_n$ be an arbitrary vector such that $\phi_i(k) = aH^T$. Then $g = a \oplus f_i(k) \in C_0$, and by (1), we have $z_i \oplus a = (s_i, u_i)G_2G_1 \oplus g \oplus v_i$, where the vector $(s_i, u_i)G_2G_1 \oplus g$ is a codeword in C_1. Hence, the adversary can recover this codeword as well as the vector v_i by applying efficient decoding algorithm for C_1 to corrupted codeword $z_i \oplus a$. Now, knowing v_i, the adversary can find the vector

$$(z_i \oplus a \oplus v_i)H^T = (s_i, u_i)G_2G_1H^T \oplus gH^T = (s_i, u_i)G_2G_1H^T.$$

Let's denote by G_1' and G_1'' submatrices contained in the first $m-l$ and in the last l rows of the matrix G_1, respectively. Using (4), we get

$$(s_i, u_i)G_2G_1 = (u_i, s_i \oplus u_iB)G_1 = u_iG_1' \oplus (s_i \oplus u_iB)G_1''.$$

Since H is a parity-check matrix of the code $C_0 = \{uG_1' \oplus uBG_1'' : u \in V_{m-l}\}$, we have $G_1'H^T = BG_1''H^T$. Thus

$$(s_i, u_i)G_2G_1H^T = (u_iG_1' \oplus (s_i \oplus u_iB)G_1'')H^T = s_iG_1''H^T.$$

So, the adversary can find s_i from the known vector $(s_i, u_i)G_2G_1H^T$ by solving the system of linear equations $x\,G_1''H^T = s_iG_1''H^T$ with respect to the unknown $x \in V_l$.

Let us remark that this system has a unique solution (equal to s_i). Indeed, in the converse case there exists a non-zero vector $x \in V_l$ such that $x\,G_1''H^T = 0$. But then $x\,G_1'' \in C_0$, and hence, there exists $u \in V_{m-l}$ such that $x\,G_1'' = uG_1' \oplus uBG_1''$. Therefore $u = 0$, $uB \oplus x = 0$ since rows of the matrix G_1 are linearly independent (see(4)). Thus, $x = 0$ that contradicts the above assumption.

As a result, we obtain the following *algorithm for the recovery of s_i from the known values z_i and $\phi_i(k) = f_i(k)H^T$.*

1. Find (by Gaussian elimination, for example) an arbitrary vector $a \in V_n$ such that $\phi_i(k) = aH^T$.
2. Recover the vector v_i by applying efficient decoding algorithm for the code C_1 to corrupted codeword $z_i \oplus a$.
3. Recover s_i as a unique solution $x \in V_l$ of the system of linear equations $x\,G_1''H^T = (z_i \oplus a \oplus v_i)H^T$.

The statement is proved.

Now, let us show that key recovery in the ciphertext-only attacking scenario can be reduced to solving a system of linear Boolean equations corrupted by noise with the noise level equal to $p_{d_1^{\perp}-1}$.

Statement 2. *Under condition (1) the complexity of recovering the secret key in the ciphertext-only attacking scenario is upper bounded by the complexity of solving the following system of linear equations corrupted by noise:*

$$z_ih^T = f_i(k)h^T \oplus \xi_i, \; i = 0, 1, \ldots, t, \qquad (7)$$

where $h \in C_1^{\perp}$ is an arbitrary codeword of weight $d_1^{\perp}$, $\xi_0, \xi_2, \ldots, \xi_t$ are independent random variables with the distribution law

$$\mathbf{P}(\xi_i = 1) = 1 - \mathbf{P}(\xi_i = 0) = p_{d_1^\perp - 1},\ i = 0, 1, ..., t. \quad (8)$$

Proof. Indeed, it follows from condition $h \in C_1^\perp$ that $(s_i, u_i) G_2 G_1 h^T = 0$ for any $s_i \in V_l$, $u_i \in V_{m-l}$. Thus, by (1), $z_i h^T = f_i(k) h^T \oplus v_i h^T$. Since the coordinates of random vector v_i are independent and take values 0 and 1 with probabilities $1-p$ and p, respectively, we have $\mathbf{P}(v_i h^T = 1) = 1 - \mathbf{P}(v_i h^T = 1) = p_{wt(h)-1}$ (see Lemma 9.49 in [4], for example). This completes the proof.

Note that in some cases, the secret key k can be uniquely recovered from the system (7) in time substantially smaller than the complexity of solving the LPN problem with the same number of unknowns and noise level defined by (8). For example, for the functions f_i of the form (2) the complexity of solving the system (7) depends essentially on algebraic properties of the sequence $S^i h^T$, $i = 0, 1, ..., t$. In particular, when a large collection of low-weight parity-check equations for this sequence is available, the specified system of equations can be efficiently solved by applying well-known algorithms used in fast correlation attacks (see [5], for example).

Chosen-Plaintext / Chosen-IV Attacks. Now, let us consider the attack described in [10, 11], where it is supposed that the adversary can encrypt the same message $s_i = 0$ with an (unknown) key getting the messages z_i of the form (1), $i = 0, 1, ..., t$. It is clear that the considerations stated above for the ciphertext-only attacking scenario are applicable in this case as well, with the only difference that instead of the code C_1 its subcode C_0 should be considered. In particular, the following statement holds.

Statement 3. *Under conditions (1), (2) the complexity of recovering the secret key in CPA scenario is upper bounded by the complexity of solving the following system of linear equations corrupted by noise:*

$$z_i h^T = k(S^i h^T) \oplus \xi_i,\ i = 0, 1, ..., t, \quad (9)$$

where $h \in C_0^\perp$ *is an arbitrary codeword of weight* $d_0^\perp$, $\xi_0, \xi_2, ..., \xi_t$ *are independent random variables with the distribution law*

$$\mathbf{P}(\xi_i = 1) = 1 - \mathbf{P}(\xi_i = 0) = p_{d_0^\perp - 1},\ i = 0, 1, ..., t.$$

Let us remark that in [10], in order to recover the secret key in CPA scenario a system of linear equations corrupted by noise is formed as well. This system differs from (9) and has a more complicated form. Moreover, the noise level in the right-hand side of this system is not less (but, can be greater) than $p_{d_0^\perp - 1}$.

In the case when the adversary has access to the encryption device and can choose (on his own) public parameters (for example, initialization vectors) determining the functions f_i, he can mount a more powerful attack by encrypting (for some fixed i) the same message $s_i = 0$ under the same IV. Note that such multiple encryptions don't give additional information about the key if an ordinary (non-randomized) cipher is used. But, for the cipher specified by (1) the adversary can derive the following equations:

$$z^{(j)} = (0, u^{(j)}) G_2 G_1 \oplus f_i(k) \oplus v^{(j)},\ j = 0, 1, ..., \quad (10)$$

where the unknown value $f_i(k)$ is fixed, and $u^{(0)}, v^{(0)}, u^{(1)}, v^{(1)}, ...$ are independent random vectors distributed as follows:

$$\mathbf{P}(u^{(j)} = u) = 2^{-(m-l)},\ \mathbf{P}(v^{(j)} = v) = p^{wt(v)}(1-p)^{n-wt(v)},$$
$$u \in V_{m-l},\ v \in V_n.$$

Using standard technique it is not hard to prove the following statement.

Statement 4. *Let* H *be an arbitrary parity-check matrix of the code* C_0, $d(H) = \max_{1 \le r \le n-m+l} wt(H_r)$, *where* H_r *is the* r *-th row of the matrix* H. *Then for any* $i = 0, 1, ...$, $k \in K$, *and* $\delta \in (0,1)$ *the adversary can recover the value* $\phi_i(k) = f_i(k) H^T$ *with probability at least* $1-\delta$ *in* $O(nt \log t)$ *bit operations from* $t = \lceil 1/2 \cdot (1-2p)^{-2d(H)} \ln(\delta^{-1}(n-m+l)) \rceil$ *arbitrary equations of the system (10).*

Proof. It follows from (10) that for any $r = 1, 2, ..., n-m+l$ the following equalities hold:

$$z^{(j)} H_r^T = f_i(k) H_r^T \oplus \xi_{j,r},\ j = 0, 1, ...,$$

where $\xi_{j,r}$ are independent random variables distributed as follows:

$$\mathbf{P}(\xi_{j,r} = 1) = 1 - \mathbf{P}(\xi_{j,r} = 0) = 1/2 \cdot \left(1 - (1-2p)^{wt(H_r)}\right),$$
$$j = 0, 1, ...$$

Suppose that to recover the value $f_i(k) H_r^T$ the majority rule is used, i.e., $f_i(k) H_r^T$ is set in 0 iff $\sum_{j=1}^{t} z^{(j)} H_r^T < t/2$. Then using the Chernoff bound we can estimate the error probability as follows:

$$\mathbf{P}\left(\sum_{j=1}^{t} \xi_{j,r} \ge t/2\right) = \mathbf{P}\left(t^{-1} \sum_{j=1}^{t} \xi_{j,r} - 1/2 \cdot \left(1 - (1-2p)^{wt(H_r)}\right) \ge (1-2p)^{wt(H_r)}\right) \le$$
$$\le \exp\{-2t(1-2p)^{2wt(H_r)}\} \le \exp\{-2t(1-2p)^{2d(H)}\}.$$

Thus, the probability of the event that all values $f_i(k)H_r^T$, $r=1,2,...,n-m+l$, are correct recovered is lower bounded by

$$1-(n-m+l)\exp\{-2t(1-2p)^{2d(H)}\}\geq 1-\delta,$$

where the last inequality follows from the definition of t. Finally, it is clear that the bit time complexity of recovering all values $f_i(k)H_r^T$, $r=1,2,...,n-m+l$, with the majority rule is upper bounded by $O(nt\log t)$. This completes the proof.

Note that chosen IV attacks are not considered in [9, 10, 12], but they should be taken into account in security analysis of randomized encryption schemes based on real keystream generators used in ordinary stream ciphers.

4. Conclusion. The computational security of randomized stream ciphers specified by (1) significantly depends on the choice of matrices G_1, G_2, and functions f_i, which are determined by the employed keystream generator, and can be much less than it is claimed in [10 – 12]. In particular, some of specified ciphers are vulnerable even to ciphertext-only attacks.

The influence of the keystream generator is demonstrated through the fact that systems of linear equations with corruptions formed to recover the key can have very dedicated form and can be solved considerably faster than the LPN problem with the same number of unknowns and noise level. The last parameter depends on the dual distance of the code C_0, whose appropriate choice (for a fixed code C_1), taking into account (6), seems a non-trivial problem. (Emphasize that the design criteria for the matrix G_2 formulated in [10], pp. 11 and 15, do not guarantee the claimed level of security).

From our point of view, to increase the security of the considered stream ciphers it is desirable to refuse from error-correction coding at all and use an encryption scheme of the following form:

$$z_i=(u_i, s_i\oplus\phi(u_i))P * f_i(k),\ i=0,1,..., \quad (11)$$

where $z_i, u_i, s_i, f_i(k)$ have the same sense as above, $n=m$, P is a permutation matrix (for example, defined by a rotation by certain number of bits), $*$ is a commutative group operation on the set V_m, and $\varphi: V_{m-l}\to V_l$ is a mapping "with good cryptographic properties" such as those used in modern block ciphers. For example, we can set $a*b=(a+b)\bmod 2^m$, where arbitrary vectors $a,b\in V_m$ are identified with the corresponded numbers in the set $\{0,1,...,2^m-1\}$, $m=2l$, and $\varphi(x)=x^{2^l-2}$, $x\in\mathbf{GF}(2^l)$. Note that in [2, 3] a similar approach for design of randomized block ciphers with provable security against some known cryptographic attacks was proposed. Another possible approach is to use a keyless hash function (such as Keccak) as the function φ. Taking into account the fact that practically secure hash function simulates a random mapping (in our case from V_{m-l} to V_l) sufficiently well, the last variant looks more preferable with regard to providing adequate security of the randomized cipher. The computational security of the stream ciphers specified by (11) is a subject of future research.

REFERENCES

[1]. Alekseychuk A. N., Gryshakov S. V., (2004) "Nonlinear random coding for information transmission systems with the wire-tap", *Legal, regulatory and metrological support information security system in Ukraine*., Vol. 8, PP. 133-140.

[2]. Alekseychuk A. N., (2007) "Analytical bounds of parameters that determine the provable security of randomized block ciphers against differential cryptanalysis", *Zakhist Inform*, No 2., PP. 12-23.

[3]. Alekseychuk A. N., (2007) "Sufficient conditions for randomized block cipher-systems to be secure against commutative diagram cryptanalysis", *Data Recording, Storage and Processing*. ,Vol. 9., No 2., PP. 61-68.

[4]. Logachev O.A., Sal'nikov A.A., Yashchenko V.V., (2004) "Boolean functions in coding theory and cryptology", MCCME, Moscow.

[5]. Canteaut A. (2005), "Fast correlation attacks against stream ciphers and related open problems", *The 2005 IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security* – ITW 2005, E-Proc. (6 p.), Awaji Island, Japan.

[6]. Gilbert H., Robshaw M.J.B., Seurin Y. (2008), "How to encrypt with the LPN problem", *ICALP 2008*, Part II, Lecture Notes in Computer Science, Vol. 5126, PP. 679-690.

[7]. MacWilliams F.J., Sloane N.J.A. (1977), "The theory of error-correcting codes", North Holland, Amsterdam.

[8]. Mihaljević M.J., Imai H. (2008), "A stream ciphering approach based on wiretap channel coding", *8th Central European Conference of Cryptography*, Graz, Austria, July 2-4, E-Proc. (3 p.).

[9]. Mihaljević M.J., Imai H. (2009), "An approach for stream cipher design based on joint computing over random and secret data", *Computing*, Vol. 85, No 1-2, June 2009, PP. 153-168.

[10]. Mihaljević M.J., Oggier F., Imai H. (2010), "Homophonic coding design for communication systems employing the encoding-encryption paradigm", in arXiv:1012.5895v1 [cs.CR], 29 Dec.

[11]. Mihaljević M.J., Imai H. (2011), "Employment of homophonic coding for improvement of certain en-

cryption approaches based on the LPN problem", *Symmetric Key Encryption Workshop* – SKEW 2011, Copenhagen, Denmark, Feb. 16-17, E-Proc. (17 p.).
[12]. Mihaljević M.J., Imai H. (2011), "An information-theoretic and computational complexity security analysis of a randomized stream cipher model", *4th Western European Workshop on Research in Cryptology* – WeWoRC 2011, Weimar, Germany, July 20-22, Conf. Record PP. 21-25.
[13]. Oggier F., Mihaljević M.J. (2010), "An information-theoretic analysis of the security of communication systems employing the encoding-encryption paradigm", in arXiv:1008.0968v1 [cs.CR], 5 Aug.
[14]. Thangaraj A., Dihidar S., Calderbank A.R., McLaughlin S.W., Merolla J.-M. (2007), "Applications of LDPC codes to the wiretap channel", *IEEE Trans. Information Theory*, Vol. 53, No 8, PP. 2933-2945.
[15]. Wyner A.D. (1975), "The wire-tap channel", Bell. Systems Technical Journal, Vol. 54, PP. 1355-1387.

ЛИТЕРАТУРА

[1]. Алексейчук А.Н., Гришаков С.В., "Нелинейное случайное кодирование в системах передачи информации по каналу связи с отводом", *Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні*. – В.8. – 2004. – С. 133-140.
[2]. Алексейчук А.Н., "Аналитические границы параметров, определяющие доказуемую стойкость рандомизированных блочных шифров относительно дифференциального криптоанализа", *Захист інформації*. – №2. – 2007. – С. 12-23.
[3]. Алексейчук А.Н.., "Достаточные условия стойкости рандомизированных блочных систем шифрования относительно метода криптоанализа на основе коммутативных диаграмм", *Реєстрація, зберігання і обробка даних*. – Т.9. – №2. – 2007. – С. 61-68.
[4]. Логачев О.А., Сальников А.А., Ященко В.В., "Булевы функции в теории кодирования и криптологии", МЦНМО, Москва, 2004.
[5]. Canteaut A. (2005), "Fast correlation attacks against stream ciphers and related open problems", *The 2005 IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security* – ITW 2005, E-Proc. (6 p.), Awaji Island, Japan.
[6]. Gilbert H., Robshaw M.J.B., Seurin Y. (2008), "How to encrypt with the LPN problem", *ICALP 2008*, Part II, Lecture Notes in Computer Science, Vol. 5126, PP. 679-690.
[7]. MacWilliams F.J., Sloane N.J.A. (1977), "The theory of error-correcting codes", North Holland, Amsterdam.
[8]. Mihaljević M.J., Imai H. (2008), "A stream ciphering approach based on wiretap channel coding", *8th Central European Conference of Cryptography*, Graz, Austria, July 2-4, E-Proc. (3 p.).
[9]. Mihaljević M.J., Imai H. (2009), "An approach for stream cipher design based on joint computing over random and secret data", *Computing*, Vol. 85, No 1-2, June 2009, PP. 153-168.
[10]. Mihaljević M.J., Oggier F., Imai H. (2010), "Homophonic coding design for communication systems employing the encoding-encryption paradigm", in arXiv:1012.5895v1 [cs.CR].
[11]. Mihaljević M.J., Imai H. (2011), "Employment of homophonic coding for improvement of certain encryption approaches based on the LPN problem", *Symmetric Key Encryption Workshop* – SKEW 2011, Copenhagen, Denmark, Feb. 16-17, E-Proc. (17 p.).
[12]. Mihaljević M.J., Imai H. (2011), "An information-theoretic and computational complexity security analysis of a randomized stream cipher model", *4th Western European Workshop on Research in Cryptology* – WeWoRC 2011, Weimar, Germany, July 20-22, Conf. Record, PP. 21-25.
[13]. Oggier F., Mihaljević M.J. (2010), "An information-theoretic analysis of the security of communication systems employing the encoding-encryption paradigm", in arXiv:1008.0968v1 [cs.CR].
[14]. Thangaraj A., Dihidar S., Calderbank A.R., McLaughlin S.W., Merolla J.-M. (2007), "Applications of LDPC codes to the wiretap channel", *IEEE Trans. Information Theory*, Vol. 53, No 8, PP. 2933-2945.
[15]. Wyner A.D. (1975), "The wire-tap channel", *Bell. Systems Technical Journal*, Vol. 54, PP. 1355-1387.

О ВЫЧИСЛИТЕЛЬНОЙ СТОЙКОСТИ РАНДОМИЗИРОВАННЫХ ПОТОЧНЫХ ШИФРОВ, ПРЕДЛОЖЕННЫХ МИХАЛЕВИЧЕМ И ИМАИ

В данной статье проводится анализ вычислительной стойкости широкого класса рандомизированных поточных шифров, построенных на основе общего применения процедур шифрования, помехоустойчивого и, соответственно, специального случайного кодирования. Показано, что стойкость указанных шифров может быть значительно меньше, чем утверждают их разработчики. В отличие от подхода к анализу стойкости, используемого в предыдущих работах, предложено более простые аналитические методы, которые позволяют выяснить теоретико-кодовый смысл параметров, определяющих вычислительную стойкость этих шифров. Предложено один из возможных альтернативных способов (на основе нелинейного случайного кодирования) построения рандомизированных поточных шифров с повышенной стойкостью.

Ключевые слова: симметричная криптография, рандомизированное шифрование, поточный шифр, случайное кодирование, отводной канал, задача LPN, корреляционная атака.

ПРО ОБЧИСЛЮВАЛЬНУ СТІЙКІСТЬ РАНДОМІЗОВАНИХ ПОТОКОВИХ ШИФРІВ, ЗАПРОПОНОВАНИХ МІХАЛЕВИЧЕМ ТА ІМАЇ

У даній статті проводиться аналіз обчислювальної стійкості широкого класу рандомізованих потокових шифрів, побудованих на основі спільного застосування процедур шифрування, завадостійкого та, відповідно, спеціального випадкового кодування. Показано, що стійкість зазначених шифрів може бути значно менше, ніж стверджують їх розробники. На відміну від підходу до аналізу стійкості, що використовується у попередніх роботах, запропоновано більш прості аналітичні методи, які дозволяють з'ясувати теоретико-кодовий сенс параметрів, що визначають обчислювальну стійкість цих шифрів. Запропоновано один із можливих альтернативних способів (на основі нелінійного випадкового кодування) побудови рандомізованих потокових шифрів із підвищеною стійкістю.

Ключові слова: симетрична криптографія, рандомізоване шифрування, потоковий шифр, випадкове кодування, відвідний канал, задача LPN, кореляційна атака.

Олексійчук Антон Миколайович, доктор технічних наук, професор Інституту спеціального зв'язку та захисту інформації НТУУ «КПІ».
E-mail: alex-dtn@ukr.net.
Алексейчук Антон Николаевич, доктор технических наук, профессор Института специальной связи и защиты информации НТУУ «КПИ».
Alekseychuk Anton, Doctor of Technical Science, Professor of Institute of Special Communication and Information Security of NTUU «KPI».

Гришаков Сергій Володимирович, здобувач Інституту спеціального зв'язку та захисту інформації НТУУ «КПІ».
E-mail: gsv-crypto@mail.ru.
Гришаков Сергей Владимирович, соискатель Института специальной связи и защиты информации НТУУ «КПИ».
Gryshakov Sergey, applicant of Institute of Special Communication and Information Security of NTUU «KPI».

УДК 004.056.2:004.421.5

МЕТОД ФОРМИРОВАНИЯ ИМИТОВСТАВКИ НА ОСНОВЕ ПЕРЕСТАНОВОК

Эмиль Фауре, Валерий Швыдкий, Валентина Щерба

Для построения защищенных телекоммуникационных систем актуальной является задача контроля целостности передаваемых сообщений, который обеспечивается за счет использования процедуры имитозащиты данных. С учетом роста производительности вычислительных средств, а также совершенствования методов взлома систем защиты информации, в том числе защиты от навязывания ложных данных, возрастают требования к стойкости методов и средств имитозащиты. В работе разработана и представлена структурная схема устройства формирования случайной последовательности перестановок. На основе принципов построения данного устройства предложен метод формирования имитовставки и устройство, его реализующее. Сущность метода заключается в том, что в качестве имитовставки используется выбранная в некотором порядке часть символов перестановки большой размерности. Указанная перестановка формируется из последовательности символов сообщения, преобразованных в последовательность взаимосвязанных чисел, представленных в факториальной системе счисления. Для скрытия закона формирования имитовставки используется сменяемый ключ преобразования. Определена стойкость перестановки и сформированной из нее имитовставки при попытке взлома ключа методом «грубой силы».

Ключевые слова: *генератор перестановок, имитозащита, имитовставка, факториальная система счисления, преобразование факториального числа в перестановку, ключ преобразования.*

Введение. Непрерывное совершенствование средств вычислительной техники, их эффективное применение для взлома систем защиты информации приводит к непрерывному процессу совершенствования методов и средств защиты, включая средства и методы имитозащиты [2]. Естественным ответом на непрерывный рост производительности ЭВМ, используемых для взлома систем защиты информации, является требование столь же быстрого роста крипто- и имитостойкости систем защиты. Это обстоятельство обуславливает актуальность разработки новых методов и средств имитозащиты данных с повышенной стойкостью.

Выделение нерешенных задач. Несмотря на несомненные успехи в области разработки технологий повышения стойкости имитозащиты, любые работы, проводимые в этом направлении, представляют значительный интерес. В частности, представляют интерес работы, связанные с разработкой и исследованием новых методов и средств синтеза (случайных) последовательностей перестановок, упрощения алгоритмов их формирования (уменьшение числа и сложности операций, уменьшение объема требуемой памяти и т.п.), в том числе на основе использования факториальной системы счисления [1, 3, 4].

Использование факториальной системы счисления предусматривает представление каж-