МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

UDC 621.391.7 DOI 10.17223/20710410/45/4

ON THE CONSTRUCTION OF A SEMANTICALLY SECURE MODIFICATION OF THE MCELIECE CRYPTOSYSTEM

Y. V. Kosolapov, O. Y. Turchenko

Southern Federal University, Rostov-on-Don, Russia

E-mail: itaim@mail.ru

The security of currently used asymmetric cryptosystems is based on the problems of discrete logarithm or discrete factorization. These problems can be effectively solved using Shor's algorithm on quantum computers. An alternative to such cryptosystems can be the McEliece cryptosystem. Its security is based on the problem of decoding a general linear code. In its original form, the McEliece cryptosystem is not semantically secure, from here the problem of constructing a semantically secure cryptosystem of the McEliece type is relevant. In the paper, the goal is to construct a McEliece type cryptosystem that has the IND-CPA property. Further, one can suppose that this system can be used as base cryptosystem for building the McEliece type encryption scheme with the IND-CCA2 property and an efficient information transfer rate.

Keywords: *McEliece type cryptosystems, IND-CPA, semantic security, standart model.*

Introduction

Many public-key cryptosystems are vulnerable to attacks on ciphertext: chosen plaintext attack, chosen ciphertext attack, malleability attack. The readers are referred to [1] for detailed description of these attacks. Semantically secure cryptosystems are immune to most of these attacks. Semantic security was introduced in [2] and means that the ciphertext does not give the adversary any information about the plaintext with polynomial restrictions on adversary's computing resources. One way to build such cryptosystems is to use probabilistic encryption. For example, M. Bellare and P. Rogaway in [3] proposed the optimal asymmetric encryption padding (OAEP) modification for the widely used asymmetric RSA cryptosystem. It should be noted that the security of currently used asymmetric cryptosystems is based on the problems of discrete logarithm or discrete factorization. These problems can be effectively solved using Shor's algorithm [4] on quantum computers. An alternative to such cryptosystems can be the McEliece cryptosystem [5], whose security is based on the problem of decoding a general linear code. In its original form, the McEliece cryptosystem is not semantically secure. The problem of constructing a semantically secure cryptosystem of the McEliece type is relevant. In [1] a modification has been constructed that possesses the strongest persistence property — the indistinguishability under adaptive chosen ciphertext attack (IND-CCA2). However, this property is achieved only in the random oracle model. This model was first used in [6] and means that protocol participants have access to some theoretical function (oracle). The oracle for any unique argument produces a truly random value and if the argument repeats, the oracle repeats the corresponding output. In [7] a modification of McEliece cryptosystem is constructed that has the property of indistinguishability under chosen plaintext attack (IND-CPA) without using the random oracle model. In this case, one can say that the standard model is used. This modification was later used in [8] as a *base cryptosystem* to construct a system that has the IND-CCA2 property within the standard model. In [8] one information message is encrypted l times, which leads to a decrease in the information transfer rate by at least l times. It is important to note that l is the length of the digital signature key. To provide high security according to [9] the key length of the asymmetric cryptosystem underlying the digital signature algorithm should be at least 256 bits. From here, the rate of information transfer of the cryptosystem from [8] is essentially low. Consequently, the development of cryptosystems of the McEliece type with the IND-CCA2 property and the high information transfer rate is current of interest.

In the present paper, the goal is to construct a McEliece type cryptosystem that has the IND-CPA property. Further, using the ideas of [8], one can suppose that this system can be used as base cryptosystem for building the McEliece type encryption scheme with the IND-CCA2 property and a higher information transfer rate.

The paper has the following structure. In Section 1 2 we introduce the basic definitions. The Section 2 describes the McEliece cryptosystem [5] and its semantically secure modification [7]. Three new cryptosystems are also constructed here. Two of them are used in Section 3 to prove the semantic security of the third one. Section 4 proposes data transfer protocol using this modification.

1. Preliminaries

Let $\mathbb{F}_q$ be a Galois field of cardinality q, where q is the degree of a prime number, $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{F}_q^n$. The support of the vector $\mathbf{m}$ is the set $\operatorname{supp}(\mathbf{m}) = \{i : m_i \neq 0\}$ and the Hamming weight of this vector is a number $\operatorname{wt}(\mathbf{m}) = |\operatorname{supp}(\mathbf{m})|$. For the vector $\mathbf{m} \in \mathbb{F}_q^n$ and the ordered set $\omega \subseteq \{1, \ldots, n\}$ we consider the projection operator $\Pi_\omega : \mathbb{F}_q^n \to \mathbb{F}_q^{|\omega|}$ acting according to the rule:

$$\Pi_\omega(\mathbf{m}) = (m_{i_1}, \ldots, m_{i_{|\omega|}}), \;\; i_j \in \omega, \; j = 1, \ldots, |\omega|.$$

Let $\mathbf{x} \in \mathbb{F}_q^{n_1}$, $\mathbf{y} \in \mathbb{F}_q^{n_2}$, $\mathbf{z} \in \mathbb{F}_q^n$, $n_1 + n_2 = n$, $\omega \subset \{1, \ldots, n\}$, $|\omega| = n_1$, then $\mathbf{z} = \mathbf{x} \parallel \mathbf{y}$ will be a concatenation of the vectors $\mathbf{x}$ and $\mathbf{y}$. Denote $\mathbf{z} = \mathbf{x} \parallel_\omega \mathbf{y}$ as merging of these vectors over an ordered set ω. In other words, $\Pi_\omega(\mathbf{z}) = \mathbf{x}$ and $\Pi_{\{1,\ldots,n\}\setminus\omega}(\mathbf{z}) = \mathbf{y}$. Further we will use the standard notations for writing algorithms and experiments described in [10]. By $y \leftarrow \mathcal{A}(x_1, x_2, \ldots)$ we mean that the algorithm $\mathcal{A}$ runs with input parameters $x_1, x_2, \ldots$ and output value y. If the algorithm $\mathcal{A}$ has access to the output of the algorithm (oracle) $\mathcal{O}$ then we write $y \leftarrow \mathcal{A}^{\mathcal{O}}(x_1, x_2, \ldots)$. If S is a finite set, then $s \in_{\mathrm{R}} S$ denotes the operation of picking an element at random and uniformly from S. To denote an asymmetric encryption scheme we will use the triplet of algorithms, i.e. $\Sigma = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, where 1) $\mathcal{K}$ is a probabilistic polynomial-time key generation algorithm which takes as input a security parameter $N \in \mathbb{N}$ and outputs a public-key pk and a secret-key sk; 2) $\mathcal{E}$ is probabilistic polynomial-time encryption algorithm which receives as input a public-key pk and a message $\mathbf{m}$, and outputs a ciphertext $\mathbf{c}$. We will write $\{\mathbf{m}\}_{pk}^{\Sigma}$ as encryption of the message $\mathbf{m}$ with the key pk; 3) $\mathcal{D}$ is deterministic polynomial-time decryption algorithm which takes as input a secret-key sk and a ciphertext $\mathbf{c}$, and outputs either a message $\mathbf{m}$ or a symbol $\perp$ in the case, when ciphertext is incorrect. Decryption of the ciphertext $\mathbf{c}$ on the secret key sk we will denote $\{\mathbf{c}\}_{sk}^{\Sigma}$.

We say a function $\gamma : \mathbb{N} \to [0,1]$ is negligible in k, if $\forall c \in \mathbb{N}\ \exists k_c\ (\gamma(k) \leqslant k^{-c}$ for all $k > k_c)$.

Now we will consider the notions of the security of public key cryptosystems. The first one is the indistinguishability under chosen plaintext attack introduced in [2]. We will consider it in the same way as [8].

Let Σ be an encryption scheme and let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be an adversary. It should be noted that $\mathcal{A}$ is polynomial time if both probabilistic algorithm $\mathcal{A}_1$ and probabilistic algorithm $\mathcal{A}_2$ are polynomial time. Now one can consider the following experiment (Algorithm 1).

Algorithm 1. $\mathbf{Exp}^{\mathrm{cpa}}_{\Sigma,\mathcal{A}}$

```
$(pk, sk) \leftarrow \mathcal{K}(\mathbf{1}^N)$;
$(\mathbf{m}_0, \mathbf{m}_1, st) \leftarrow \mathcal{A}_1(pk)$;
$b \leftarrow \{0,1\}$;
$\mathbf{c} \leftarrow \{\mathbf{m}_b\}^{\Sigma}_{pk}$;
$B \leftarrow \mathcal{A}_2(\mathbf{c}, st)$.
If $B = b$, then return 1, else return 0.
```

The meaning of this experiment can be explained by an example. Let Σ be the basic RSA cryptosystem over ring $\mathbb{Z}_n$. The adversary selects two plaintexts using the algorithm $\mathcal{A}_1$ which generates messages randomly or by using some features of the cryptosystem. In the basic RSA cryptosystem the feature is the fact that $\{0\}^{\Sigma}_{pk} = 0 \in \mathbb{Z}_n$ for any pk. Let $\mathcal{A}_1$ always gives a pair $(0, a, st)$, where $a \neq 0$ and st is the whole state information obtained during the run of $\mathcal{A}_1$. For instance st contains a public key pk and generated messages $\mathbf{m}_0$, $\mathbf{m}_1$. Then the experimenter selects random coin b and encrypts $\mathbf{m}_b$. The adversary's task, given the encryption $\mathbf{c}$, is to determine which of the two plaintexts was encrypted. In the framework of this example, algorithm $\mathcal{A}_2$ can be trivial. In fact $\mathcal{A}_2$ checks whether the resulting cipher is a zero number. If it is, then $\mathcal{A}_2$ outputs 0 (corresponds to zero plain text), otherwise 1 (corresponds to plain text a).

The advantage of the adversary $\mathcal{A}$ is determined by the value

$$\mathbf{Adv}^{\mathrm{cpa}}_{\Sigma,\mathcal{A}}(N) = \left| \mathsf{P}[\mathbf{Exp}^{\mathrm{cpa}}_{\Sigma,\mathcal{A}} = 1] - \frac{1}{2} \right|,$$

where $\mathsf{P}[A]$ denotes probability of the event A. It is said that the cryptosystem Σ has the property IND-CPA if for any polynomial algorithm $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ the advantage of $\mathbf{Adv}^{\mathrm{cpa}}_{\Sigma,\mathcal{A}}(N)$ is a negligible function in N.

Now let the adversary $\mathcal{A}^D = (\mathcal{A}^{\mathcal{D}}_1, \mathcal{A}^{\mathcal{D}}_2)$ has access to the decryption oracle $\mathcal{D}$. By $\mathcal{A}^{\mathcal{D}\{\cdot\}}_i$ we mean that adversary $\mathcal{A}^{\mathcal{D}}_i$ has a polynomial number of queries to the oracle $\mathcal{D}$. Let us consider the following experiment (Algorithm 2).

Algorithm 2. $\mathbf{Exp}^{\mathrm{cca2}}_{\Sigma,\mathcal{A}}$

```
$(pk, sk) \leftarrow \mathcal{K}(\mathbf{1}^N)$;
$(\mathbf{m}_0, \mathbf{m}_1, st) \leftarrow \mathcal{A}^{\mathcal{D}\{\cdot\}}_1(pk)$;
$b \leftarrow \{0,1\}$;
$\mathbf{c}^* \leftarrow \{\mathbf{m}_b\}^{\Sigma}_{pk}$;
$B \leftarrow \mathcal{A}^{\mathcal{D}\{\cdot\}}_2(\mathbf{c}^*, st)$, and $\mathcal{D}\{\mathbf{c}^*\} = \perp$;
If $B = b$, then return 1, otherwise 0.
```

The principal difference from the previous experiment is that the algorithms $\mathcal{A}_1$ and $\mathcal{A}_2$ have access to decryption oracle. The decryption oracle takes as input a ciphertext and for a polynomial time outputs the corresponding plain text. The only limitation is that this oracle can not be requested by the cipher text produced by the experimenter on step 4 ($\mathcal{D}\{\mathbf{c}^*\} =\perp$). In [11], a practical attack on the RSA standard PKCS #1 was presented (the SSL protocol used that standard at that time), which was based on the idea of decryption oracle.

The advantage of adversary $\mathcal{A}^D$ is

$$\mathbf{Adv}^{\text{cca2}}_{\Sigma,\mathcal{A}}(N) = \left|\mathsf{P}[\mathbf{Exp}^{\text{cca2}}_{\Sigma,\mathcal{A}} = 1] - \frac{1}{2}\right|.$$

It is said that the cryptosystem Σ has the property IND-CCA2 if for any polynomial algorithm $\mathcal{A}^D$ an advantage $\mathbf{Adv}^{\text{cca2}}_{\Sigma,\mathcal{A}}(N)$ is negligible function in N.

Further we need some notions from [12, p. 22–26]. Let X_0 and X_1 be finite random variables with the set of values $\mathbf{D}$. Then the statistical distance is the function

$$\delta(X_0, X_1) = \frac{1}{2}\sum_{d\in\mathbf{D}} \big|\,\mathsf{P}[X_0 = d] - \mathsf{P}[X_1 = d]\big|.$$

Let $\mathbb{A}$ be a class of polynomial-time algorithms, which take a cipher text $\mathbf{c}$ and some state information st as input and output one bit. For example, within the framework of the experiment $\mathbf{Exp}^{\text{cpa}}_{\Sigma,\mathcal{A}}$ algorithm $\mathcal{A}_2$ belongs to this class.

Then we will say that ciphertexts of two different cryptosystems $\Sigma_1 = (\mathcal{K}, \mathcal{E}_1, \mathcal{D}_1)$ and $\Sigma_2 = (\mathcal{K}, \mathcal{E}_2, \mathcal{D}_2)$ are indistinguishable by the class of polynomial algorithms $\mathbb{A}$ if for any information message $\mathbf{m}$ and for all $\mathbf{A} \in \mathbb{A}$

$$\delta(\mathbf{A}(\{\mathbf{m}\}^{\Sigma_1}_{pk_1}, st_1), \mathbf{A}(\{\mathbf{m}\}^{\Sigma_2}_{pk_2}, st_2))$$

is a negligible function in N, where pk_i is generated by $\mathcal{K}(\mathbf{1}^N)$. It is not difficult to verify that for all $\mathbf{A} \in \mathbb{A}$

$$\delta(\mathbf{A}(\{\mathbf{m}\}^{\Sigma_1}_{pk_1}, st_1), \mathbf{A}(\{\mathbf{m}\}^{\Sigma_2}_{pk_2}, st_2)) = \big|\,\mathsf{P}[\mathbf{A}(\{\mathbf{m}\}^{\Sigma_1}_{pk_1}, st_1) = 0] - \mathsf{P}[\mathbf{A}(\{\mathbf{m}\}^{\Sigma_2}_{pk_2}, st_2) = 0]\big|.$$

Lemma 1. Let $\Sigma_1 = (\mathcal{K}, \mathcal{E}_1, \mathcal{D}_1)$ and $\Sigma_2 = (\mathcal{K}, \mathcal{E}_2, \mathcal{D}_2)$ are cryptosystems, Σ_1 has the IND-CPA property. If ciphertexts of two different cryptosystems are indistinguishable by the class of polynomial algorithms $\mathbb{A}$, then Σ_2 has the IND-CPA property.

Proof. Suppose that there is an adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ such that $\mathbf{Adv}^{\text{cpa}}_{\Sigma_2,\mathcal{A}}(N)$ is a function ψ that is not negligible in N. Now we construct the adversary algorithm $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ on the basis of $\mathcal{A}$ and estimate $\mathbf{Adv}^{\text{cpa}}_{\Sigma_1,\mathcal{B}}(N)$. Let pk_1 is public key generated by $\mathcal{K}$. The algorithm $\mathcal{B}_1$ takes as input pk_1 and generates public key pk_2 using $\mathcal{K}$. Then $\mathcal{B}_1$ calls the algorithm $\mathcal{A}_1(pk_2)$ and outputs a triplet $(\mathbf{m}_0, \mathbf{m}_1, st_1)$. Thus, in spite of different public keys, the outputs of $\mathcal{B}_1(pk_1)$ and $\mathcal{A}_1(pk_2)$ will be identical. The algorithm $\mathcal{B}_2$ simply calls the $\mathcal{A}_2$ algorithm from its input. Since the experiments $\mathbf{Exp}^{\text{cpa}}_{\Sigma_1,\mathcal{B}}$ and $\mathbf{Exp}^{\text{cpa}}_{\Sigma_2,\mathcal{A}}$ differ on fourth step, the outputs of the algorithms $\mathcal{B}_2$ and $\mathcal{A}_2$ may differ. Consider the statistical distance between their outputs. By the condition of the lemma, the ciphers are indistinguishable by the class of polynomial algorithms $\mathbb{A}$. Since $\mathcal{A}_2$ belongs to this class, then $\big|\,\mathsf{P}[\mathcal{A}_2(\{\mathbf{m}\}^{\Sigma_1}_{pk_1}, st_1)=0] - \mathsf{P}[\mathcal{A}_2(\{\mathbf{m}\}^{\Sigma_2}_{pk_2}, st_2) = 0]\big| = \eta$, where η is a negligible function in N. Because of $\mathcal{B}_2$ simply calls $\mathcal{A}_2$ we have

$$|\,\mathsf{P}[\mathcal{B}_2(\{\mathbf{m}\}^{\Sigma_1}_{pk_1}, st_1) = 0] - \mathsf{P}[\mathcal{A}_2(\{\mathbf{m}\}^{\Sigma_2}_{pk_2}, st_2) = 0]| = \eta.$$

It follows that $\mathbf{Adv}^{\mathrm{cpa}}_{\Sigma_1,\mathcal{B}}(N) = \psi \pm \eta$, as $\mathbf{Adv}^{\mathrm{cpa}}_{\Sigma_1,\mathcal{B}}(N)$ is directly related to the output of $\mathcal{B}_2$. But $\psi \pm \eta$ is not a negligible in N. This contradicts the fact that Σ_1 has the IND-CPA property. ■

2. McEliece type cryptosystems

Consider the McEliece cryptosystem $\mathrm{McE}(C)$ on the linear $[n,k,d]$-code $C(\subseteq \mathbb{F}_q^n)$, where n is the length, k is the code dimension, and d is the minimum code distance. Let G be the generating matrix of the code C, $t = \lfloor (d-1)/2 \rfloor$. A secret key sk is a pair (S,P), where S is a non-singular $(k \times k)$-matrix over the field $\mathbb{F}_q$, and P is a permutation $(n \times n)$-matrix. A public key pk is a pair $(\widetilde{G} = SGP, t)$. Encryption of a message $\mathbf{x} \in \mathbb{F}_q^k$ is performed according to the rule

$$\{\mathbf{x}\}^{\mathrm{McE}}_{pk} = \mathbf{x}\widetilde{G} + \mathbf{e} = \mathbf{y},\ \mathrm{wt}(\mathbf{e}) \leqslant t.$$

To decrypt the ciphertext $\mathbf{y}$ one should use an effective decoder $\mathrm{Dec}_C : \mathbb{F}_q^n \to \mathbb{F}_q^k$ of the code C and the secret key sk:

$$\{\mathbf{y}\}^{\mathrm{McE}}_{sk} = \mathrm{Dec}_C(\mathbf{y}P^{-1})S^{-1}. \tag{1}$$

For the same code C, we consider the modification $\mathrm{McE}_l(C)$ of the McEliece type cryptosystem described in [7], where encryption rule has the form

$$\{\mathbf{x}\}^{\mathrm{McE}_l}_{pk} = \{\mathbf{x} \parallel \mathbf{v}\}^{\mathrm{McE}}_{pk} = \mathbf{y},\ \mathbf{x} \in \mathbb{F}_q^l,\ \mathbf{v} \in_R \mathbb{F}_q^{k-l}. \tag{2}$$

To decrypt the ciphertext $\mathbf{y}$, it is enough to apply the rule (1) and discard the last $k-l$ symbols:

$$\{\mathbf{y}\}^{\mathrm{McE}_l}_{sk} = \{\mathbf{y}\}^{\mathrm{McE}}_{sk} (I_l \parallel O_{n-l})^\top,$$

where I_l is the unit $(l \times l)$ matrix, O_{k-l} is the zero $(k-l \times k-l)$ matrix, and $A^\top$ is the transposed matrix A.

On the basis of the cryptosystem $\mathrm{McE}_l(C)$ we construct a new cryptosystem $2\mathrm{McE}_l(C)$, in which the message of length l is encrypted twice according to the rule (2):

$$\{\mathbf{x}\}^{2\mathrm{McE}_l}_{pk} = \{\mathbf{x}\}^{\mathrm{McE}_l}_{pk} \parallel \{\mathbf{x}\}^{\mathrm{McE}_l}_{pk} = \mathbf{y},\ \mathbf{x} \in \mathbb{F}_q^l.$$

Then the decryption rule can be written in the form:

$$\{\mathbf{y}\}^{2\mathrm{McE}_l}_{sk} = \left\{\mathbf{y}\,(I_n \parallel O_n)^\top\right\}^{\mathrm{McE}_l}_{sk}.$$

Consider a subset $\mathcal{G}_l$ of permutations group $\mathcal{S}_k$ acting on the elements of the set $\{1,\ldots,k\}$ such that for any $\pi \in \mathcal{G}_l$ the condition $\pi(1) < \ldots < \pi(l)$ is satisfied. The set $\{\pi(1),\ldots,\pi(l)\}$ is denoted by ω_π. Note that $|\mathcal{G}_l| = \mathrm{C}_k^l (k-l)!$, since only C_k^l subsets of cardinality l are in the set of k elements, and for each such subset ω there is a class $\mathcal{G}(\omega) \subseteq \mathcal{S}_k$ permutations with cardinality $|\mathcal{G}(\omega)| = (k-l)!$. With every permutation π from $\mathcal{G}_l$ we associate a permutation $(k \times k)$-matrix R_π. Consider the cryptosystem $\omega 2\mathrm{McE}'_l$ with the encryption rule

$$\{\mathbf{x}\}^{\omega 2\mathrm{McE}'_l}_{pk} = \{(\mathbf{x} \parallel \mathbf{v}_1)R_\pi\}^{\mathrm{McE}}_{pk} \parallel \{(\mathbf{x} \parallel \mathbf{v}_2)R_\pi\}^{\mathrm{McE}}_{pk} = \mathbf{y}, \tag{3}$$

where $\mathbf{x} \in \mathbb{F}_q^l, \mathbf{v}_i \in_{\mathrm{R}} \mathbb{F}_q^{k-l}, i = 1,2, \pi \in_R \mathcal{G}_l$. For decryption, in addition to the secret key sk, the recipient needs to know the matrix R_π. Then the decryption rule takes the form

$$\{\mathbf{y}\}^{\omega 2\mathrm{McE}'_l}_{sk,R_\pi} = (\{\mathbf{y}\,(I_n \parallel O_n)^\top\}^{\mathrm{McE}}_{sk} \cdot R_\pi^{-1})(I_l \parallel O_{k-l})^\top.$$

Finally, we construct a cryptosystem $\omega 2\mathrm{McE}_l$ based on previous one with the following restriction: $\mathrm{supp}(\mathbf{v}_1 - \mathbf{v}_2) = \{1, \ldots, k\} \setminus \omega_\pi$. Then, for decryption, the recipient does not need the matrix R_π. To find ω, it suffices to compute the vector

$$\mathbf{z} = \{\mathbf{y}\,(I_n \parallel O_n)^\top\}_{sk}^{\mathrm{McE}} - \{\mathbf{y}\,(O_n \parallel I_n)^\top\}_{sk}^{\mathrm{McE}}$$

and find its support $\mathrm{supp}\,(\mathbf{z})$. Then the decryption rule takes the form

$$\{\mathbf{y}\}_{sk}^{\omega 2\mathrm{McE}_l} = (\mathbf{z} \cdot R_{\pi'}^{-1})(I_l \parallel O_{k-l})^\top, \pi' \in \mathcal{G}_l(\omega), \omega = \mathrm{supp}\,(\mathbf{z})\,.$$

3. Semantic security of McEliece type cryptosystems

3.1. Security assumptions

Let $\mathrm{McE}(C)$ be the basic McEliece cryptosystem with security parameter N. The security of $\mathrm{McE}(C)$ is based on the problem of decoding a random linear code [5]. Note that, if there is no polynomial algorithm capable of distinguishing the $(k \times n)$-matrix of the public key of the $\mathrm{McE}(C)$ cryptosystem from a random $(k \times n)$-matrix with non-negligible probability in N, then the cryptosystem $\mathrm{McE}_l(C)$ has the IND-CPA property [7].

Further we will use two additional assumptions.

Assumption 1. There is no polynomial algorithm that can distinguish two random noisy codewords of the code C from random vectors with a non-negligible probability in security parameter N.

The assumption is based on the fact that at present there are no such polynomial algorithms. For example, recent algorithms [13 – 15] that solve the given problem are not polynomial.

Assumption 2. There is no polynomial algorithm that takes as input ciphertext $\mathbf{c}$ of the $\mathrm{McE}(C)$ and the number $l \in \mathbb{N}$, and outputs 0 if $\mathbf{c}$ corresponds to an information message of a weight less than l and outputs 1 if $\mathbf{c}$ corresponds to an information message of weight l with non-negligible distinguishing advantage in the N.

3.2. IND-CPA security of $2\mathrm{McE}_l(C)$

It is easy to verify that the cryptosystem $\mathrm{McE}(C)$ is not IND-CPA-secure for an arbitrary $[n, k, d]$-code C. At the same time, the cryptosystem $\mathrm{McE}_l(C)$ on the Goppa code C is IND-CPA-secure [7].

Let us consider the matrix $\widetilde{G}$ of the public key of the cryptosystem $\mathrm{McE}_l(C)$ in the form

$$\widetilde{G} = \begin{pmatrix} \widetilde{G}_1 \\ \widetilde{G}_2 \end{pmatrix},$$

where $\widetilde{G}_1$ is $(l \times n)$-matrix and $\widetilde{G}_2$ is $(kl \times n)$-matrix. To prove IND-CPA-security of cryptosystem $2\mathrm{McE}_l(C)$ consider the algorithm $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2)$ and following experiment (Algorithm 3).

It is important to note that the algorithm $\mathcal{D}_2$ takes a decision only by two vectors and does not accumulate vectors.

Suppose that there exists a polynomial $g(N)$, a polynomial algorithm $\mathcal{D}' = (\mathcal{D}'_1, \mathcal{D}'_2)$ and an infinite subsequence of natural numbers $(N_1, N_2, \ldots)$ such that for all $i = 1, 2, \ldots$ the following inequality holds:

$$\mathsf{P}[\mathbf{Exp}_{\widetilde{G},t,\mathcal{D}'}^{\mathrm{dif}_1}(N_i) = 1] \geqslant \frac{1}{2} + \frac{1}{q(N_i)}. \tag{4}$$

Algorithm 3. $\mathbf{Exp}^{\mathrm{dif}_1}_{\widetilde{G},t,\mathcal{D}}$

```
$\mathbf{m}_0 \leftarrow \mathcal{D}_1(N)$;
$b \leftarrow \{0,1\}$.
If $b = 1$, then $\mathbf{c} = \{\mathbf{m}_0\}^{2\mathrm{McE}_l}_{(\widetilde{G},t)}$, otherwise $\mathbf{c} \in_R \mathbb{F}_q^{2n}$.
$B \leftarrow \mathcal{D}_2(\mathbf{c}, \mathbf{m}_0)$;
If $B = b$, return 1, otherwise 0.
```

In other words, the algorithm $\mathcal{D}'$ with a non-negligible probability distinguishes one pair of ciphertexts corresponding to one information message from a pair of random vectors. Let's construct one more algorithm $\mathcal{WD}$ and experiment $\mathbf{Exp}^{\mathrm{dif}_2}_{\widetilde{G}_2,t,\mathcal{WD}}$ (Algorithm 4).

Algorithm 4. $\mathbf{Exp}^{\mathrm{dif}_2}_{\widetilde{G}_2,t,\mathcal{WD}}$

```
$b \leftarrow \{0,1\}$.
If $b = 0$, then $\mathbf{y}_1, \mathbf{y}_2 \in_{\mathrm{R}} \mathbb{F}_q^n$, otherwise $\mathbf{y}_i = \mathbf{r}_i\widetilde{G}_2 + \mathbf{e}_i$, $\mathbf{r}_i \in_{\mathrm{R}} \mathbb{F}_q^{k-l}$, $\mathrm{wt}(\mathbf{e}_i) \leqslant t$, $i = 1, 2$.
$B \leftarrow \mathcal{WD}(\mathbf{y}_1, \mathbf{y}_2, \widetilde{G}_2, t)$.
If $B = b$, then return 1, else 0.
```

In the experiment $\mathbf{Exp}^{\mathrm{dif}_2}_{\widetilde{G}_2,t,\mathcal{WD}}$, given algorithm $\mathcal{WD}$ distinguishes two random noisy codewords of the code with the generator matrix $\widetilde{G}_2$ from random vectors. From here, using $\mathcal{D}'$ one can construct polynomial algorithm $\mathcal{WD}'$ to solve this distinguishing problem (Algorithm 5).

Algorithm 5. $\mathcal{WD}'(\mathbf{y}_1, \mathbf{y}_2, \widetilde{G}_2, t)$

```
$\mathbf{m}_0 \leftarrow \mathcal{D}'_1(N)$;
$\mathbf{c}' = (\mathbf{m}_0\widetilde{G}_1 + \mathbf{y}_1) \parallel (\mathbf{m}_0\widetilde{G}_1 + \mathbf{y}_2)$.
Return $\mathcal{D}'_2(\mathbf{c}', \mathbf{m}_0)$.
```

Given (4), we get : $\mathsf{P}[\mathbf{Exp}^{\mathrm{dif}_2}_{\widetilde{G}_2,t,\mathcal{WD}'}] \geqslant \dfrac{1}{2} + \dfrac{1}{q(N_i)}$. But it contradicts the assumption 1.

Hence we obtain that for any polynomial algorithm $\mathcal{D}'$ and any polynomial $q(N)$, the following inequality holds:

$$\left|\mathsf{P}[\mathbf{Exp}^{\mathrm{dif}_1}_{\widetilde{G},t,\mathcal{D}'}(N) = 1] - \frac{1}{2}\right| < \frac{1}{q(N)}. \tag{5}$$

Note that for the experiment $\mathbf{Exp}^{\mathrm{dif}_3}_{\widetilde{G},t,\mathcal{M}}$ (Algorithm 6) the probability of occurrence of 1 is also differs from 1/2 by a negligibly small function. Otherwise, based on corresponding algorithm, one can construct an algorithm $\mathcal{WD}''$ with not negligible $|\,\mathsf{P}[\mathbf{Exp}^{\mathrm{dif}_2}_{\widetilde{G}_2,t,\mathcal{WD}''}(N) = 1] - 1/2|$ in N.

Hence it follows that there is no polynomial algorithm $\mathcal{Q}$ that distinguishes the ciphertext $\{\mathbf{m}_0\}^{2\mathrm{McE}_l}_{(\widetilde{G},t)}$ from $[\{\mathbf{m}_1\}^{\mathrm{McE}_l}_{(\widetilde{G},t)} \parallel \{\mathbf{m}_2\}^{\mathrm{McE}_l}_{(\widetilde{G},t)}]$ with a probability that is not negligible greater than 1/2 for any $\mathbf{m}_0$, $\mathbf{m}_1$, $\mathbf{m}_2$. Otherwise, we can construct the polynomial algorithm $\widetilde{D} = (\widetilde{D}_1, \widetilde{D}_2)$ (see Algorithm 7) for the experiment $\mathbf{Exp}^{\mathrm{dif}_1}_{\widetilde{G},t,\mathcal{M}}$, which with a non-negligible probability would distinguish a pair of ciphertexts from a pair of random vectors, which contradicts the assumption.

Algorithm 6. $\mathbf{Exp}^{\mathrm{dif}_3}_{\widetilde{G},t,\mathcal{M}}$

```
$\mathbf{m}_1, \mathbf{m}_2 \leftarrow \mathcal{M}_1(N)$;
$b \leftarrow \{0,1\}$.
If $b = 1$, then $\mathbf{c} = [\{\mathbf{m}_1\}^{\mathrm{McE}_l}_{(\widetilde{G},t)} \parallel \{\mathbf{m}_2\}^{\mathrm{McE}_l}_{(\widetilde{G},t)}]$, otherwise $\mathbf{c} \in_{\mathrm{R}} \mathbb{F}_q^{2n}$.
$b' \leftarrow \mathcal{M}_2(\mathbf{c}, \mathbf{m}_1, \mathbf{m}_2)$.
If $B = b$, return 1, otherwise 0.
```

Algorithm 7. $\widetilde{D}_2(\mathbf{c}_0, \mathbf{m}_0)$

```
$\mathbf{m}_1, \mathbf{m}_2 \leftarrow \mathcal{M}_1(N)$;
$\mathbf{c}_1 = \{\mathbf{m}_1\}^{\mathrm{McE}_l}_{pk} \parallel \{\mathbf{m}_2\}^{\mathrm{McE}_l}_{pk} + \mathbf{c}$;
$v \leftarrow \{0,1\}$.
Return $\mathcal{Q}(\mathbf{c}_v, \mathbf{m}_0, \mathbf{m}_1 + \mathbf{m}_0, \mathbf{m}_2 + \mathbf{m}_0)$.
```

Theorem 1. If the cryptosystem $\mathrm{McE}_l(C)$ has the IND-CPA property, then the cryptosystem $2\mathrm{McE}_l(C)$ also has this property .

Proof. Suppose that the cryptosystem $2\mathrm{McE}_l(C)$ does not have the IND-CPA property. Then there exists a polynomial algorithm (adversary) $\mathcal{A}' = (\mathcal{A}'_1, \mathcal{A}'_2)$, the polynomial $p(N)$ and an infinite subsequence of natural numbers $(N_1, N_2, \ldots)$ such that for all $i = 1, 2, \ldots$ the following inequality holds:

$$\mathbf{Adv}^{\mathrm{cpa}}_{2\mathrm{McE}_l(C),\mathcal{A}'}(N_i) \geqslant \frac{1}{p(N_i)}. \tag{6}$$

On the basis of the algorithm $\mathcal{A}'$ we construct the algorithm $\mathcal{A}'' = (\mathcal{A}'_1, \mathcal{A}''_2)$ for the attack on the cryptosystem McE_l. The algorithm $\mathcal{A}''_2$ takes as input the ciphertext $\mathbf{c} = \{\mathbf{m}_b\}^{\mathrm{McE}_l}_{pk}$ of the cryptosystem $\mathrm{McE}_l(C)$ and two messages $\mathbf{m}_0$, $\mathbf{m}_1$; the algorithm $\mathcal{A}''_2$ randomly picks up the value v from $\{0,1\}$ and returns the result $\mathcal{A}'_2(\mathbf{c} \parallel \{\mathbf{m}_v\}^{\mathrm{McE}_l(C)}_{pk})$. Then

$$\mathsf{P}[\mathbf{Exp}^{\mathrm{cpa}}_{\mathrm{McE}_l,\mathcal{A}''} = 1] = \frac{1}{2}\,\mathsf{P}[\mathcal{A}'_2(\{\mathbf{M}\}^{\mathrm{McE}_l(C)}_{pk} \parallel \{\mathbf{M}'\}^{\mathrm{McE}_l(C)}_{pk}) = 1 | \mathbf{M} = \mathbf{M}'] +$$
$$+ \frac{1}{2}\,\mathsf{P}[\mathcal{A}'_2(\{\mathbf{M}\}^{\mathrm{McE}_l(C)}_{pk} \parallel \{\mathbf{M}'\}^{\mathrm{McE}_l(C)}_{pk}) = 1 | \mathbf{M} \neq \mathbf{M}'].$$

From (6) we get $|\,\mathsf{P}[\mathcal{A}'_2(\{\mathbf{M}\}^{\mathrm{McE}_l(C)}_{pk} \parallel \{\mathbf{M}'\}^{\mathrm{McE}_l(C)}_{pk}) = 1 | \mathbf{M} = \mathbf{M}'] - 1/2| \geqslant 1/p(N_i)$, and from the explanation , which comes after the (5), we have

$$\left| \mathsf{P}[\mathcal{A}'_2(\{\mathbf{M}\}^{\mathrm{McE}_l(C)}_{pk} \parallel \{\mathbf{M}'\}^{\mathrm{McE}_l(C)}_{pk}) = 1 | \mathbf{M} \neq \mathbf{M}'] - \frac{1}{2} \right| < \frac{1}{q(N)}.$$

In this way,

$$\left| \mathsf{P}[\mathbf{Exp}^{\mathrm{cpa}}_{\mathrm{McE}_l,\mathcal{A}''} = 1] - \frac{1}{2} \right| > \frac{1}{p(N_i)} \pm \phi(N),$$

where $\phi(N)$ is a negligibly small function. Since $\dfrac{1}{p(N_i)} \pm \phi(N)$ is not a negligibly small function, we have obtained that the cryptosystem $\mathrm{McE}_l(C)$ does not have the property $\mathrm{IND-CPA}$, which contradicts the condition. ■

3.3. IND-CPA property for $\omega 2\mathrm{McE}'_l(C)$

Lemma 2. If the cryptosystem $2\mathrm{McE}_l(C)$ has the IND-CPA property, then the cryptosystem $\omega 2\mathrm{McE}'_l(C)$ also has this property.

Proof. The encryption rule (3) can be rewritten as

$$\{\mathbf{x}\}_{pk}^{\omega 2\mathrm{McE}'l} = ((\mathbf{x} \parallel \mathbf{v}_1)R_\pi \widetilde{G} \oplus \mathbf{e}_1) \parallel ((\mathbf{x} \parallel \mathbf{v}_2)R_\pi \widetilde{G} \oplus \mathbf{e}_2).$$

Denote $\widetilde{G}' = R_\pi \widetilde{G}$. Then we get

$$\{\mathbf{x}\}_{pk}^{\omega 2\mathrm{McE}'l} = ((\mathbf{x} \parallel \mathbf{v}_1\widetilde{G}' \oplus \mathbf{e}_1) \parallel ((\mathbf{x} \parallel \mathbf{v}_2)\widetilde{G}' \oplus \mathbf{e}_2) = \{\mathbf{x}\}_{pk'}^{\mathrm{McE}_l} \parallel \{\mathbf{x}\}_{pk'}^{\mathrm{McE}_l} = \{\mathbf{x}\}_{pk'}^{2\mathrm{McE}_l},$$

where $pk' = (\widetilde{G}', t)$. Thus by construction $\omega 2\mathrm{McE}'_l(C)$ is the same as $2\mathrm{McE}_l(C)$ but with different pair (pk, sk). From here $\omega 2\mathrm{McE}'_l(C)$ also has the IND-CPA property. ■

Note, that adversary doesn't know the relationships between (pk, sk) and (pk', sk'). From here adding a permutation in the $2\mathrm{McE}_l(C)$ cryptosystem with the help of the set ω can only increase the security.

3.4. IND – CPA - property for $\omega 2\mathrm{McE}_l(C)$

Theorem 2. The cryptosystem $\omega 2\mathrm{McE}_l(C)$ has the IND-CPA property if the cryptosystem $\omega 2\mathrm{McE}'_l(C)$ has this property.

Proof. For the proof it is sufficiently to show that the ciphertexts of cryptosystems $\omega 2\mathrm{McE}_l(C)$ and $\omega 2\mathrm{McE}'_l(C)$, corresponding to one information message, are indistinguishable by the class of algorithms $\mathbb{A}$. We fix an arbitrary $\mathbf{m}$ and consider the ciphertexts of cryptosystems $\omega 2\mathrm{McE}_l(C)$ and $\omega 2\mathrm{McE}'_l(C)$ as a system of the form:

$$\{\mathbf{m}\}_{pk}^{\omega 2\mathrm{McE}_l} = X \parallel Y, \quad \begin{cases} X = \mathbf{m}\widetilde{G}^1_\omega \oplus \mathbf{r}_1\widetilde{G}^2_\omega \oplus \mathbf{e}_1, \\ Y = \mathbf{m}\widetilde{G}^1_\omega \oplus (\mathbf{1} \oplus \mathbf{r}_1)\widetilde{G}^2_\omega \oplus \mathbf{e}_2, \end{cases}$$

$$\{\mathbf{m}\}_{pk}^{\omega 2\mathrm{McE}'l} = X \parallel Y', \quad \begin{cases} X = \mathbf{m}_1\widetilde{G}^1_\omega \oplus \mathbf{r}_1\widetilde{G}^2_\omega \oplus \mathbf{e}'_1, \\ Y' = \mathbf{m}_1\widetilde{G}^1_\omega \oplus \mathbf{r}_2\widetilde{G}^2_\omega \oplus \mathbf{e}'_2. \end{cases}$$

Denote $\mathbf{r}'_2 = \mathbf{r}_1 \oplus \mathbf{r}_2$. Then the systems can be rewritten:

$$\{\mathbf{m}\}_{pk}^{\omega 2\mathrm{McE}_l} = X \parallel Y, \quad \begin{cases} X = \mathbf{m}\widetilde{G}^1_\omega \oplus \mathbf{r}_1\widetilde{G}^2_\omega \oplus \mathbf{e}_1, \\ Y = \mathbf{m}\widetilde{G}^1_\omega \oplus \mathbf{r}_1\widetilde{G}^2_\omega \oplus \mathbf{1}\widetilde{G}^2_\omega \oplus \mathbf{e}_2, \end{cases}$$

$$\{\mathbf{m}\}_{pk}^{\omega 2\mathrm{McE}'l} = X \parallel Y', \quad \begin{cases} X = \mathbf{m}\widetilde{G}^1_\omega \oplus \mathbf{r}_1\widetilde{G}^2_\omega \oplus \mathbf{e}'_1, \\ Y' = \mathbf{m}\widetilde{G}^1_\omega \oplus \mathbf{r}_1\widetilde{G}^2_\omega \oplus \mathbf{r}'_2\widetilde{G}^2_\omega \oplus \mathbf{e}'_2. \end{cases}$$

Now we consider the last parts of Y and Y': $LP = \mathbf{1}\widetilde{G}^2_\omega \oplus \mathbf{e}_2$ and $LP' = \mathbf{r}'_2\widetilde{G}^2_\omega \oplus \mathbf{e}'_2$. Denote Z as $Y = Z \oplus LP$ and Z' as $Y' = Z' \oplus LP'$. Since (5) we get that the rest $X \parallel Z'$ does not provide any information about LP. One should note that $X \parallel Z' = X \parallel Z$. From here $X \parallel Z'$ does not provide any information about LP'. Consequently, to distinguish the ciphertexts one should distinguish LP and LP'. A vector LP of the form $\mathbf{1}\widetilde{G}^2_\omega \oplus \mathbf{e}_1$ can be rewritten as $(\mathbf{0} \parallel \mathbf{1})R_\pi \widetilde{G} \oplus \mathbf{e}_1$. Thus, for a random choice of ω, LP is a ciphertext of basic McEliece cryptosystem corresponding to a random information message with a fixed weight l. The vector $LP' = (\mathbf{r}_1 \oplus \mathbf{r}_2)\widetilde{G}^2_\omega \oplus \mathbf{e}_1$ can similarly be rewritten as $(\mathbf{0} \parallel \mathbf{r}_1 \oplus \oplus \mathbf{r}_2)R_\pi \widetilde{G} \oplus \mathbf{e}_1$ and is also a ciphertext of the basic McEliece cryptosystem, but corresponding to a random information message of arbitrary weight not exceeding l. By Assumption 2, algorithm for distinguishing vectors of this kind does not exist. Hence the ciphertexts of cryptosystems $\omega 2\mathrm{McE}_l(C)$ and $\omega 2\mathrm{McE}'_l(C)$, corresponding to one information message, are indistinguishable by the class of algorithms $\mathbb{A}$. ■

4. Implementation of ω2McE

We suppose a possible implementation of ω2McE to modify k-repetition scheme [8]. The idea of k-repetition scheme is to encrypt information message k-times using INC-CPA-secure cryptosystem Σ. Encryption of k-repetition scheme has the form $\{\mathbf{m}\}^{\Sigma}_{pk_1} \parallel \parallel \{\mathbf{m}\}^{\Sigma}_{pk_2} \parallel \ldots \parallel \{\mathbf{m}\}^{\Sigma}_{pk_k}$. Note that to encryption requires k unique key pairs. We suggest use ω2McE in k-repetition scheme with some modifications. The idea of our modification is to encrypt $k/2$ information messages using only one set ω. So encryption will take the form $\{\mathbf{m_1}\}^{\omega 2\mathrm{McE}}_{pk_1} \parallel \{\mathbf{m_2}\}^{\omega 2\mathrm{McE}}_{pk_2} \parallel \ldots \parallel \{\mathbf{m_{k/2}}\}^{\omega 2\mathrm{McE}}_{pk_{k/2}}$. In fact, it also requires to encrypt k-times. Let us remind that k is the length of signature key and should be more than 512. However, our construction transmits $k/2$ information messages. From here, with this approach, the data transfer rate will increase by $k/2$ times.

REFERENCES

1. *Kobara K. and Imai H.* Semantically secure McEliece public-key cryptosystems — conversions for McEliece PKC. LNCS, 2001, vol. 1992, pp. 19–35.
2. *Goldwasser S. and Micali S.* Probabilistic encryption. J. Computer and System Sciences, 1984, vol. 38, no. 2, pp. 270–299.
3. *Bellare M. and Rogaway P.* Optimal asymmetric encryption — how to encrypt with RSA. Advances in Cryptology — EUROCRYPT'94, Springer Verlag, 1995, pp. 92–111.
4. *Shor P.* Algorithms for quantum computation: discrete logarithms and factoring. Proc. 35th Ann. Symp. FCS, Santa Fe, USA, IEEE Publ., 1994, pp. 124–134.
5. *McEliece R. J.* A public-key cryptosystem based on algebraic coding theory. DSN Progress Report, 1978, vol. 42, no. 44, pp. 114–116.
6. *Bellare M. and Rogaway P.* Random oracles are practical: A paradigm for designing efficient protocols. CCS '93 Proc. 1st ACM conf. CCS'93, N.Y., ACM, 1993, pp. 62–73.
7. *Nojima R., Imai H., Kobara K., and Morozov K.* Semantic security for the McEliece cryptosystem without random oracles. Designs, Codes and Cryptography, 2008, vol. 49, no. 1–3, pp. 289–305.
8. *Dottling N., Dowsley R., Muller-Quade J., and Nascimento C. A. A.* A CCA2 secure variant of the McEliece cryptosystem. IEEE Trans. Inform. Theory, 2012, vol. 58, no. 10, pp. 6672–6680.
9. *Lenstra A. K. and Verheul E. R.* Selecting cryptographic key sizes. J. Cryptology, 2001, vol. 14, no. 4, pp. 255–293.
10. *Bellare M., Desai A., Pointcheval D., and Rogaway P.* Relations among notions of security for public-key encryption schemes. Advances in Cryptology — CRYPTO'98, LNCS, 1998, vol. 1462, pp. 26–45.
11. *Bleichenbacher D.* Chosen ciphertext attacks against protocols based on the RSA encryption standard PKCS#1. Advances in Cryptology — CRYPTO'98, LNCS, 1998, vol. 1462, pp. 255–293.
12. *Cramer R., Damgard I., and Nielsen J. B.* Secure Multiparty Computation and Secret Sharing. Cambridge, Cambridge University Press, 2015. 373 p.
13. *Kosolapov Y. V. and Turchenko O. Y.* Primenenie odnogo metoda raspoznavaniya koda dlya kanala s podslushivaniem [Application of one method of linear code recognition to the wire-tap channel]. Prikladnaya Diskretnaya Matematika, 2017, no. 35, pp. 76–88. (in Russian)

14. *Chabot C.* Recognition of a code in a noisy environment. Proc. IEEE ISIT, Nice, France, 2007, pp. 2211–2215.

15. *Yardi A. D. and Vijayakumaran S.* Detecting linear block codes in noise using the GLRT. IEEE Intern. Conf. Communications, Budapest, Hungary, 2013, pp. 4895–4899.