ISSN 0032-9460, Problems of Information Transmission, 2022, Vol. 58, No. 2, pp. 184–201. © Pleiades Publishing, Inc., 2022.
Russian Text © The Author(s), 2022, published in Problemy Peredachi Informatsii, 2022, Vol. 58, No. 2, pp. 92–111.

INFORMATION PROTECTION

On New Problems in Asymmetric Cryptography Based on Error-Resistant Coding

V. V. Zyablov[a,*], F. I. Ivanov[a,b,], E. A. Krouk[b,***], and V. R. Sidorenko[a,c,****]**

[a] *Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia*
[b] *Higher School of Economics—National Research University, Moscow, Russia*
[c] *Technische Universität München, Munich, Germany*
e-mail: *zyablov@iitp.ru, **fivanov@hse.ru, ***ekrouk@hse.ru, ****vladimir.sidorenko@tum.de

Received September 30, 2020; revised April 14, 2022; accepted April 16, 2022

Abstract—We consider the problem of constructing a cryptosystem with a public key based on error-resistant coding. At present, this type of cryptosystems is believed to be able to resist the advent of quantum computers and can be considered as a method of post-quantum cryptography. The main drawback of a code-based cryptosystem is a great length of the public key. Most papers devoted to reducing the cryptosystem key length consisted in replacing the Goppa codes used in the original cryptosystem with some other codes with a requirement that the system remains secure against attacks by a quantum computer. Here we propose another approach to the key length reduction that is stated as a task of a simple description of an error set which has either errors of weights greater than half the minimum distance or errors that cannot be corrected without an additional secret knowledge. If a code structure allows to give such a description of an error set, then the complexity of most attacks (for instance, information-set decoding) significantly increases.

Key words: McEliece cryptosystem, information-set decoding, generalized Reed–Solomon code, post-quantum cryptography.

DOI: 10.1134/S0032946022020077

1. INTRODUCTION

Methods of error-resistant coding have been used in cryptography for a long time. Using them, one of the first public key cryptosystems [1] and one of the first digital signature systems [2] were developed. However, unlike the algebraic cryptosystems based on the factoring problem [3] and discrete logarithm calculation [4], coding cryptosystems are almost not used in practice. Although coding cryptosystems have gain over algebraic ones in ciphering/deciphering time [5], their usage is largely restricted by a number of objective and subjective factors.

Firstly, algebraic systems were developed earlier than coding systems, and they immediately passed a number of security tests. This fact, in the lack of evidence of the resistance of public key cryptosystems, is a certain security guarantee.

Secondly, for the first public key cryptosystems (the McEliece cryptosystem) a public key usually has a larger length as compared to algebraic systems (for instance, the RSA systems).

Further investigation of coding cryptosystems allowed to significantly reduce the public key size [6, 7], and the development of new methods for solving the factoring problem [8] made the public key size increase so much that public key sizes for different systems became commensurable [9]. Nevertheless, algebraic cryptosystems remain the main instrument for cryptographic security nowadays.

This situation began to change in recent years due to the emergence of the notion of post-quantum cryptography [10] and the development of a number of fields where cryptographic methods can be used.

Active investigations in the field of creating the so-called quantum computer, which performs computations at a quantum level, led to the construction of algorithms focused on this computer. One of the main achievements in quantum algorithm theory is Shor's development of a polynomial algorithm for solving the factoring problem with a quantum computer [11]. The invention of this algorithm means that after creation of a powerful enough quantum computer, security systems based on algebraic cryptosystems (forming the vast majority) will be compromised. Thereby, the concept of post-quantum cryptography arose, i.e., cryptography that would remain secure after the development of quantum computers. The decoding complexity for linear codes underlying the coding cryptosystems is *NP*-hard [12], and it seems that quantum computers will not be capable of solving this task in polynomial time.

The modern practice in sensor networks, cloud computing, and a number of other areas in info-communication technologies puts forth the task of creating the so-called "light cryptography," i.e., cryptographic algorithms providing a sufficient level of security while using devices with limited computational resources [13]. For the purposes of light cryptography, coding cryptosystems turn out to be more promising [14] than algebraic ones. Coding cryptosystems require less operations and use linear algebra operations, the implementation of which is superior to arithmetic operations.

All of the above determined a new raise of interest in coding cryptography and probably a new applied phase in its development.

As was mentioned above, there were a number of attempts to overcome the main disadvantage of coding cryptosystems, a large length of a public key. The main idea of these improvements was to replace the binary Goppa code used in the original McEliece cryptosystem with some other code with a specific structure that allows reducing the public key size. For instance, in [6] Goppa codes were replaced with subcodes of quasicyclic generalized Reed–Solomon codes. This made it possible to obtain a cryptosystem with key size from 6000 to 11000 bits and with security ranging from 2^{80} to 2^{107}. In [15] the authors suggested to use quasi-cyclic moderate-density parity-check (QC-MDPC) codes. This led to significantly reducing the public key to 0.6KB, which makes cryptosystem based on these codes practical. A similar cryptosystem based on quasi-cyclic low-density parity-check (QC-LDPC) codes was suggested in [16].

The main drawback of the replacement of Goppa codes with QC-LDPC or QC-MDPC codes is that their usable iterative decoding algorithms do not guarantee correcting errors of given weight t even for relatively small weights. Moreover, practical QC-MDPC and QC-LDPC coding constructions usually have small minimum distance (about few dozens for codes with rate $R = 1/2$ and length of several thousand).

In this paper we propose a new approach to the choice of a code to be used as a part of a coding cryptosystem: instead of the problem of a compact description of a public key on account of the code structure, we set forth the problem of choosing a triple $(C_0, \mathcal{E}, \varphi)$, where C_0 is a secret (n, k, d) code, $\mathcal{E}$ a set of errors, and φ a polynomial complexity procedure which maps the set of input errors to the set of errors corrected by the code. Thus, the problem is posed of describing a set of errors (not necessarily those with a small weight) correctable by C_0. The structure of this code should be hidden. Classical attacks on the ciphertext (for example, information-set decoding [17]) face a much harder task than correction of errors of weight less than $\frac{d-1}{2}$, since they have to correct errors of weight greater than $\frac{d-1}{2}$, and this is how the reducing of the public key length is achieved. This, in turn, allows to switch to significantly shorter codes with preserving the necessary attack complexity.

2. McELIECE CRYPTOSYSTEM

The first coding cryptosystem was the McEliece cryptosystem proposed in 1978 in [1]. As a public key, a binary $k \times n$ matrix $\boldsymbol{G}$ was used which can be given as a matrix product

$$\boldsymbol{G} = \boldsymbol{S}\boldsymbol{G}_0\boldsymbol{P}, \tag{1}$$

where $\boldsymbol{S}$ is a nonsingular $k \times k$ matrix, $\boldsymbol{G}_0$ is a generator matrix of a binary (n, k, d) code C_0 for which there is a "simple" (usually polynomial time) decoding algorithm ξ for errors with multiplicity less than half the code distance $t = \dfrac{d-1}{2}$, and $\boldsymbol{P}$ is an $n \times n$ permutation matrix. We should note that the matrix $\boldsymbol{S}\boldsymbol{G}_0$ generates the same set of codewords as $\boldsymbol{G}_0$, i.e., the code C_0.

The ciphertext for a message $\boldsymbol{x}$ in the McEliece system is generated as follows:

1. Generate a random vector $\boldsymbol{e}$ of length n from the set $\mathcal{E}_t$ of the vectors of weight t;
2. Calculate a ciphertext
$$\boldsymbol{y} = \boldsymbol{x}\boldsymbol{G} + \boldsymbol{e}. \tag{2}$$

It is assumed that a legal receiver knows the matrices $\boldsymbol{S}$, $\boldsymbol{G}$, and $\boldsymbol{P}$ in the product (1), which are a private key of the cryptosystem. In this case the receiver finds an encrypted message in the following way:

1. Multiplies $\boldsymbol{y}$ by $\boldsymbol{P}^{-1}$:
$$\boldsymbol{y}\boldsymbol{P}^{-1} = \boldsymbol{x}\boldsymbol{S}\boldsymbol{G}_0 + \boldsymbol{e}\boldsymbol{P}^{-1};$$
2. Decodes the obtained vector using the code C_0 with generator matrix $\boldsymbol{S}\boldsymbol{G}_0$. Since the vector $\boldsymbol{e}\boldsymbol{P}^{-1}$ is of weight t, the result of the decoding is $\boldsymbol{x}\boldsymbol{S}$;
3. Calculates $\boldsymbol{x}$ as $\boldsymbol{x} = \boldsymbol{x}\boldsymbol{S}\boldsymbol{S}^{-1}$.

The idea of the McEliece system is that the error vector $\boldsymbol{e}\boldsymbol{P}^{-1}$ does not change its weight after applying $\boldsymbol{P}^{-1}$ to the ciphertext $\boldsymbol{y}$; i.e., it lies in the same vector set as the original error vector $\boldsymbol{e}$. That is why a legal user does not need to know an error vector for deciphering, since any error vector with weight less than t is decoded with code C_0 using the same algorithm.

The public key of the McEliece cryptosystem is the matrix $\boldsymbol{G}$. Actually, the security of the described system is based on the decoding complexity for a linear code C with an arbitrary structure. After the matrix transformation given by (1), the algebraic structure of the code matrix with a simple decoding algorithm ξ is hidden. Right multiplication by $\boldsymbol{P}$ maps the original code to an equivalent code C to which the simple decoding algorithm ξ cannot be applied.

3. ATTACKS ON CODING CRYPTOSYSTEMS

In this section, we consider a classification of attacks on coding cryptosystems (regardless of the choice of a specific code for a cryptosystem).

There are two main types of attacks on a coding cryptosystem: decoding attacks and structural attacks. Their main difference is that a decoding attack is oriented at extracting an encrypted message $\boldsymbol{x}$ from a ciphertext $\boldsymbol{y} = \boldsymbol{x}\boldsymbol{G} + \boldsymbol{e}$ via decoding $\boldsymbol{y}$ with the use of some specially selected algorithm. A structural attack aims at recovering a secret key $(\boldsymbol{S}, \boldsymbol{G}_0, \boldsymbol{P})$ from a public key $\boldsymbol{G} = \boldsymbol{S}\boldsymbol{G}_0\boldsymbol{P}$ using some available information about the structure of the code C_0 with generator matrix $\boldsymbol{G}_0$. We should note that it is not necessary to find the original factorization $\boldsymbol{G} = \boldsymbol{S}\boldsymbol{G}_0\boldsymbol{P}$. Usually, it suffices to find some factorization $(\boldsymbol{S}', \boldsymbol{G}_0', \boldsymbol{P}')$ such that $\boldsymbol{G} = \boldsymbol{S}'\boldsymbol{G}_0'\boldsymbol{P}'$, as was done in the Sidelnikov–Shestakov attack on the cryptosystem based on generalized Reed–Solomon codes [18]. In this paper, we will not dwell on the analysis of structural attacks in relation to the cryptosystem proposed by us, since we focus in more detail on decoding attacks. Moreover, it cannot be

guaranteed that there does not exist a structural attack for a given cryptosystem. It is known that the majority of successful attacks (of polynomial complexity) on various coding cryptosystems are precisely structural attacks.

Decoding attacks can be classified as follows:

1. Brute force attack on the information vector set; it is performed by exhaustive search over all possible vectors $\boldsymbol{x}$ until the equality $\mathrm{wt}(\boldsymbol{xG}-\boldsymbol{y})=t$ is obtained. The number of attempts for this attack for an (n,k) code C can be estimated from above as $2^{\min(k,n-k)}$;
2. Brute force attack on the error vector set; it is performed by exhaustive search over all possible vectors $\boldsymbol{e}$ until the equality $\mathrm{wt}((\boldsymbol{y}-\boldsymbol{e})\boldsymbol{H}^{\mathrm{T}})=0$ is obtained, where $\boldsymbol{H}$ is a parity-check matrix corresponding to the public generator matrix $\boldsymbol{G}$. The complexity of this attack depends on the cardinality of the set of errors added at the encryption stage. If the cryptosystem is based on a binary (n,k) code correcting t errors that are randomly added during the encryption stage, then the average number of attempts for this attack is $\binom{n}{t}$;
3. Attack based on searching for error free information sets. A detailed description of this attack is given in the next section.

We should note that the goal of any attack is to find a code with a simple decoding in the class of equivalent codes if it is known that a simple decoding ξ exists for at least one of the codes. The difference between the two types of attacks, structural and decoding, is that in the case of decoding attacks general decoding methods for linear codes with an arbitrary structure are used, while successful application of a structural attack allows to use a polynomial time decoder ξ.

3.1. Information Set Decoding

The minimum distance decoding problem for a code with an arbitrary structure, as was already noted in the introduction, is *NP*-hard [12]. Decoding of errors of multiplicity up to t (up to half the code distance) is relatively simpler but also has exponential complexity, although no proof of its *NP*-hardness has been presented so far. In any case, no polynomial algorithms for solving this problem are currently known.

Next, we describe the information set decoding (ISD) algorithm, on which most "promising" attacks on the McEliece cryptosystem are based.

The purpose of the ISD algorithms is to recover a message $\boldsymbol{x}$ from a given vector $\boldsymbol{y}=\boldsymbol{xG}+\boldsymbol{e}$, where $\boldsymbol{G}$ is a generator matrix of an (n,k) code C with minimum distance $d=2t+1$ and $\mathrm{wt}(\boldsymbol{e})\le t$.

Let $\mathcal{I}$ be a subset of size k of the coordinate set $[n]:=\{1,2,\dots,n\}$ such that $\mathcal{I}$ is the information set of a code C and $\boldsymbol{G}_{\mathcal{I}}$ is a submatrix of $\boldsymbol{G}$ consisting of columns with indices from $\mathcal{I}$. Similarly, let $\boldsymbol{e}_{\mathcal{I}}$ be a vector consisting of the coordinates of $\boldsymbol{e}$ with indices from $\mathcal{I}$.

The ISD algorithm works as follows:

1. Select a random information set $\mathcal{I}\subset\{1,2,\dots,n\}$;
2. If $\mathrm{wt}(\boldsymbol{y}-\boldsymbol{y}_{\mathcal{I}}\boldsymbol{G}_{\mathcal{I}}^{-1}\boldsymbol{G})\le t$, then $\boldsymbol{y}_{\mathcal{I}}$ does not contain any error, which means $\mathrm{wt}(\boldsymbol{e}_{\mathcal{I}})=0$. Then $\boldsymbol{u}=\boldsymbol{y}_{\mathcal{I}}\boldsymbol{G}_{\mathcal{I}}^{-1}$. Otherwise, return to Step 1.

It is easily seen that the probability P_k that a given information set does not contain errors is estimated from below as

$$P_k\le\frac{\binom{n-t}{k}}{\binom{n}{k}}=\frac{\binom{n-k}{t}}{\binom{n}{t}}. \tag{3}$$

This means that the average number of attempts to find an error-free information set is not greater than $\binom{n}{t}\Big/\binom{n-k}{t}$, which is significantly less than for the exhaustive search over all error

vectors. Thus, to reach the necessary security range of a McEliece cryptosystem, codes with larger lengths n are required. In particular, McEliece suggested using the $(1024, 524, 101)$ Goppa code correcting $t = 50$ errors, for which the public key length is 536 576.

The ISD attack had already been mentioned in [1] and was further developed in numerous papers (see, e.g., [19] and references therein). There are different interpretations and modifications of the original ISD algorithm. Several different improvements have been proposed, such as improvements based on the generalized birthday paradox. In [20], it was shown that the complexity exponent for information set decoding is $\widetilde{\mathcal{O}}(2^{0.0494n})$, which is presently the best known result.

3.2. Complexity of Decoding Attacks on the McEliece Cryptosystem

We provide an example of evaluating the attack complexity. (*Here and in what follows, the attack complexity means the average number of elementary operations required to find a message* $\boldsymbol{x}$ *given a ciphertext* $\boldsymbol{y}$). For the classical McEliece cryptosystem based on the $(1024, 524, 101)$ Goppa code correcting $t = 50$ errors, we have the following:

- The complexity of the attack based on the exhaustive search over the information vector set is $2^{n-k}(n-k)k = 2^{518}$;
- The complexity of the attack based on the exhaustive search over the error vector set is $k(n-k)\binom{n}{t} = 524 \cdot 500 \cdot \binom{1024}{50} \approx 2^{302}$;
- The complexity of the attack based on searching for an error-free information set can be estimated from above as

$$\frac{\binom{n}{t}}{\binom{n-k}{t}} k(n-k) = \frac{\binom{1024}{50}}{\binom{500}{50}} 524 \cdot 500 \approx 2^{72}.$$

As is noted above, the complexity of the last attack turns out to be the least. Thus, the security level (*by the security level of a cryptosystem we understand the smallest complexity among all known attacks*) of the classical McEliece cryptosystem can be estimated by the minimum complexity among the considered attacks, which is 2^{72}.

Next, we discuss ways to increase the complexity of this attack by modifying the set of errors.

4. CONSTRUCTION OF A SET OF ERRORS CORRECTED BY A CODE

All decoding algorithms described in the literature cited above are somehow or other based on the fact that they correct "light" errors with weights much less than the codeword length. However, any linear code is capable of correcting a considerable number of errors with large weight. Let a linear code C_0 of length n be defined over a field $\mathbb{F}$. We divide the space $\mathbb{F}^n$ into cosets of C_0. It is clear that C_0 can correct only one error vector from each coset, but this can be any vector from this class. This means that in expression (2) any (not necessarily "light") vectors can be used as a random masking vector (error vector $\boldsymbol{e}$) if they belong to different cosets. In this case, the complexity of the information set decoding algorithm increases significantly (it grows exponentially with t), and to ensure the required system security, a code of a much smaller size can be used.

However, another problem arises here, which was easily solved (and was not even regarded as a problem) in the original McEliece system. How to generate an error set that a legal user will decode using the code C_0 with a generator matrix $\boldsymbol{G}_0$? In fact, we must specify a set of errors $\mathcal{E}$ from which the error vector used in encryption (2) is randomly selected.

First, we formulate the properties that this set should possess. We denote by $\mathcal{E}_0$ the set of error vectors $\boldsymbol{e}$ corrected by the code C_0 using the polynomial time decoding algorithm ξ and denote

by $\varphi(C_0)$ some invertible transformation of the code (of its generator matrix, i.e., of the basis). As a result of this transformation, a code $C = \varphi(C_0)$ is generated. Let, in addition, ω be the required security level of the cryptosystem.

Now we can formulate requirements for the set $\mathcal{E}$:

(a) An algorithm V for generating a random vector $\boldsymbol{e} \in \mathcal{E}$ exists and is of polynomial complexity;
(b) There is an invertible linear transformation $\varphi(C_0)$ of polynomial complexity that maps the code C_0 to a code C;
(c) $\varphi^{-1}(\mathcal{E}) \subset \mathcal{E}_0$, and $|\varphi^{-1}(\mathcal{E})| \geq \omega$;
(d) For the code C_0, the algorithm ξ for correcting errors from $\mathcal{E}$ exists and is of polynomial complexity;
(e) The decoding complexity for errors from the set $\mathcal{E}$ in the code C is not less than ω; in particular, $|\mathcal{E}| \geq \omega$, i.e., the set cardinality must prevent brute force search over all its elements aimed at breaking the cryptosystem.

Taking into account the introduced notation, the proposed generalized encryption scheme $\mathbb{S}$ can be described as follows:

- Public key in the proposed system: the generator matrix $\boldsymbol{G} = \varphi(\boldsymbol{G}_0)$ and the algorithm V;
- Private key: the inverse transformation φ^{-1}, the matrix $\boldsymbol{G}_0$, and the decoder ξ;
- Encryption algorithm:
  1. By using the algorithm V, select a random vector $\boldsymbol{e} \in \mathcal{E}$;
  2. Given $\boldsymbol{x}$, calculate the ciphertext
$$\boldsymbol{y} = \boldsymbol{x}\boldsymbol{G} + \boldsymbol{e}; \tag{4}$$
- Decryption algorithm:
  1. Calculate $\boldsymbol{y}' = \varphi^{-1}(\boldsymbol{y}) = \boldsymbol{x}\varphi^{-1}(\boldsymbol{G}) + \varphi^{-1}(\boldsymbol{e}) = \boldsymbol{x}\boldsymbol{G}_0 + \boldsymbol{e}'$, where $\boldsymbol{e}' \in \mathcal{E}_0$;
  2. Using the algorithm ξ, find $\boldsymbol{x}$.

Although at first sight the decoding problem (4) completely coincides with problem (1), it should be more difficult due to the fact that correcting errors from $\mathcal{E}$ by a code C with an arbitrary structure is a more difficult task than correction of errors of small multiplicity. This directly follows from the fact that for given n and k, the probability P_k in (3) is a monotone decreasing function of t, i.e., $t < t' \leq \frac{n}{2}$ implies $P_k(t) \gg P_k(t')$.

Evaluating the efficiency of a coding cryptosystem. The security of a cryptosystem defined by the encryption algorithm (4), taking into account conditions (a)–(e) imposed on the set $\mathcal{E}$, is given under a direct attack (i.e., an attack based on decoding of errors from $\mathcal{E}$ in the code C) by the parameter ω. If errors from $\mathcal{E}$ are not coset leaders (the "lightest" ones in a coset), then decoding them for the code C (without knowing the inverse transformation φ^{-1}) can only be performed by exhaustive search over either the codewords from C or the error set $\mathcal{E}$. Thus,

$$\omega = \min\{2^k, 2^{n-k}, |\mathcal{E}|\}.$$

By the construction, $|\mathcal{E}| \leq 2^{n-k}$. Taking into account the fact that while decoding by exhaustive search over the codewords from C, the decoding complexity does not depend on $\mathcal{E}$, it is natural to call the cryptosystem (4) optimal if $|\mathcal{E}| = 2^{n-k}$ and use the quantity

$$\tau = \frac{\log_2 |\mathcal{E}|}{n-k} \tag{5}$$

to evaluate how "completely" the correcting properties of the code are exploited.

The formulated criterion of "completeness" of a coding cryptosystem is not sufficient nor even the most important from the point of view of practical use of the cryptosystem. It does not take into

account the size of the public key of the cryptosystem (i.e., of the code used in it), which is usually discussed as the main disadvantage of code-based cryptosystems. Therefore, along with the defined parameter, when evaluating the cryptosystem we also consider the size of the code used. Further, in addition to the brute force attack, there are a number of non-brute-force attacks, the complexity of which must be taken into account when evaluating a cryptosystem. In addition, a small value of the parameter τ indicates that "resources" of the code underlying the cryptosystem are not fully used, which means that it is potentially possible to improve the cryptosystem by expanding the set of errors which will subsequently be corrected by the decoder. On the contrary, values τ close to one allow us to state that the code underlying the cryptosystem is used quite efficiently, which means that further expansion of the set $\mathcal{E}_0$ by adding "heavy" error vectors to it will only slightly reduce the key length.

For example, for the classical McEliece cryptosystem based on the $(1024, 524)$ Goppa code, τ_{ME} is estimated as

$$\tau_{\mathrm{ME}} \geq \frac{\log_2 \binom{1024}{50}}{500} \approx 0.5681,$$

and the public key length is $1024 \cdot 524 = 536576$.

5. CRYPTOSYSTEM BASED ON THE BINARY IMAGE OF A GENERALIZED REED–SOLOMON CODE

5.1. Generalized Reed–Solomon Code and Its Binary Image

Here and in what follows we assume that the considered codes are defined over the field $\mathbb{F}_q$, $q = 2^m$, $m > 0$.

The construction of a set $\mathcal{E}$ fulfilling conditions (a)–(e) from Section 4 for a random code C is a difficult task. Nevertheless, the binary image of a generalized Reed–Solomon code (RS code) over $\mathbb{F}_q$ has a polynomial algorithm V for constructing such sets of sufficiently large cardinality.

First, we recall the definition of the generalized RS code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$.

Definition 1. Fix a finite field $\mathbb{F}_q$. Select nonzero elements $v_1, \ldots, v_n \in \mathbb{F}_q$ and different elements $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$. Let $\boldsymbol{v} = (v_1, \ldots, v_n)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$. For any $0 \leq k \leq n$, define the generalized Reed–Solomon code as

$$GRS_{n,k}(\alpha, \boldsymbol{v}) = \{(v_1 f(\alpha_1), v_2 f(\alpha_2), \ldots, v_n f(\alpha_n)) \mid f(x) \in F_k[x]\},$$

where $F_k[x]$ stands for the set of polynomials $f(x)$ over $\mathbb{F}_q$ with degrees at most $k-1$.

It is known that along with an ordinary RS code, $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ is also a maximum distance separable (MDS) code, i.e., has minimum distance $d = n - k + 1$. The main reason why this paper considers the generalized RS code is that for given n and k the cardinality of the set of different $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ codes is significantly greater than the number of different Reed–Solomon codes, which prevents a structural attack on the cryptosystem based on $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$. The generator matrix of $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ is denoted by $\boldsymbol{G}'$. This matrix is defined over the field $\mathbb{F}_q$ and is of size $k \times n$.

Let us fix a basis $\mathbb{F}_q/\mathbb{F}_2$. We consider a binary image of $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$, i.e., a code with codewords obtained from the codewords of $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ by replacing symbols over $\mathbb{F}_q$ with their binary images. Finally, we obtain a binary (nm, km) code with a $km \times nm$ generator matrix $\boldsymbol{G}'_b$. We denote this code by C_b.

The code C_b in the binary Hamming metric has minimum distance not less than that of $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$, and it is capable of correcting any error burst provided that the burst covers no more than t symbols of a received codeword if they are considered as elements of $\mathbb{F}_q$. If only the

burst length is limited to $1 \le \ell_i \le m$, but there is no limitation on the starting and ending positions of an error burst, then the number of guaranteed correctable error bursts is $\left\lfloor \frac{n-k}{4} \right\rfloor = \left\lfloor \frac{t}{2} \right\rfloor$. This follows from the fact that any $\left\lfloor \frac{t}{2} \right\rfloor$ error bursts of length $1 \le \ell_i \le m$ cannot corrupt more than t symbols over $\mathbb{F}_q$, and therefore a received vector will be corrected by the $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ code when making the inverse transformation $\mathbb{F}_2^{mn} \mapsto \mathbb{F}_q^n$.

Before proceeding to the description of the cryptosystem, we introduce the notion of a synchronous and nonsynchronous error burst.

Definition 2. An error burst of length $1 \le \ell_i \le m$ is synchronous if for its starting position i there exists an $r \in \mathbb{N} \cup 0$ such that $i \ge mr + 1$ and at the same time $i + \ell_i - 1 \le m(r+1)$, i.e., all nonzero elements of the burst are localized in one subvector $\boldsymbol{e}_{r+1}$ of the vector $\boldsymbol{e} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$. Otherwise, the burst is said to be nonsynchronous.

5.2. Basic Description of the Cryptosystem Protocol

Now we present a description of a public-key cryptosystem based on the binary image of a generalized Reed–Solomon code. In fact, this section presents a high-level description of the proposed cryptosystem, and a presentation of its individual components will be given in subsequent sections of the paper.

The public generator matrix of the cryptosystem has the form

$$\boldsymbol{G} = \boldsymbol{S}\boldsymbol{G}'_b\boldsymbol{Q}, \tag{6}$$

where $\boldsymbol{G}'_b$ is a secret binary generator matrix of a code C_b, and $\boldsymbol{S}$ is an arbitrary nonsingular binary matrix of size $mk \times mk$. The binary $mn \times mn$ matrix $\boldsymbol{Q}$ is selected according to Theorem 1 from the matrix set described in Section 5.3.

Now we describe the procedures for key generation, encryption, and decryption.

- Generation of secret and public keys:
  1. Select a generator matrix $\boldsymbol{G}'$ of the code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ and construct its binary image $\boldsymbol{G}'_b$;
  2. Construct a random nonsingular binary matrix $\boldsymbol{S}$ of size $mk \times mk$;
  3. According to Theorem 1, construct an $mn \times mn$ matrix $\boldsymbol{Q}$ and select the corresponding class of error vectors $\mathcal{V}_i$;
  4. Calculate a public generator matrix $\boldsymbol{G} = \boldsymbol{S}\boldsymbol{G}'_b\boldsymbol{Q}$;
  5. The public cryptosystem key is $(\boldsymbol{G}, \mathcal{V}_i)$;
  6. The private cryptosystem key is the set $(\boldsymbol{Q}, \boldsymbol{G}'_b, \boldsymbol{S})$.
- Encryption of the open text $\boldsymbol{x} \in \mathbb{F}_2^{km}$ is processed as follows:
  1. Select a random vector $\boldsymbol{e} \in \mathcal{V}_i \subset \mathbb{F}_2^{mn}$ consistent with the matrix $\boldsymbol{Q}$, so that the error vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ can be corrected by the code with generator matrix $\boldsymbol{G}'_b$;
  2. Calculate the ciphertext $\boldsymbol{y} \in \mathbb{F}_2^{mn}$:
  $$\boldsymbol{y} = \boldsymbol{x}\boldsymbol{G} + \boldsymbol{e}.$$
- Decryption of the vector $\boldsymbol{y} \in \mathbb{F}_2^{mn}$ is processed as follows:
  1. Calculate the product of $\boldsymbol{y}$ and $\boldsymbol{Q}^{-1}$:
  $$\boldsymbol{y}\boldsymbol{Q}^{-1} = \boldsymbol{x}\boldsymbol{S}\boldsymbol{G}'_b + \boldsymbol{e}\boldsymbol{Q}^{-1};$$
  2. The vector $\boldsymbol{y}\boldsymbol{Q}^{-1}$ is transformed into a q-vector and then decoded by the t-error-correcting code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$, whence $\boldsymbol{x}' = \boldsymbol{x}\boldsymbol{S} \in \mathbb{F}_q^k$, a vector of length k over $\mathbb{F}_q$, is found;
  3. The vector $\boldsymbol{x}' \in \mathbb{F}_q^k$ is mapped to a binary vector $\boldsymbol{x}''$;

4. The encrypted message $\boldsymbol{x}$ can be found as

$$\boldsymbol{x} = \boldsymbol{x}''\boldsymbol{S}^{-1}.$$

As is noted above, the main requirement that the pair $(\boldsymbol{Q}, \boldsymbol{e})$ must satisfy is that the vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ should be correctable by the code with generator matrix $\boldsymbol{G}'_b$, i.e., should contain no more than t synchronous or $\left\lfloor \frac{t}{2} \right\rfloor$ nonsynchronous error bursts of length up to m. Below we will show how the structures of the vectors $\boldsymbol{e}$ and matrices $\boldsymbol{Q}$ should be consistent so that $\boldsymbol{e}\boldsymbol{Q}^{-1} \in \mathcal{E}_0$, where $\mathcal{E}_0$ is the set of errors correctable by the $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ code.

5.3. Selection of $(\boldsymbol{Q}, \boldsymbol{e})$ in the Proposed Cryptosystem

Let us show how the pair $(\boldsymbol{Q}, \boldsymbol{e})$ should be chosen so that the error vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ is correctable by the binary image of the generalized Reed–Solomon code (the q-ary representation of the vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ is correctable by the code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$); in this case we say that *the vector $\boldsymbol{e}$ is consistent with the matrix $\boldsymbol{Q}$*.

Let us introduce the following notation: $\mathcal{A}(\boldsymbol{Q}, \boldsymbol{e})$ means that a vector $\boldsymbol{e}$ is consistent with a matrix $\boldsymbol{Q}$, i.e., $\boldsymbol{e}\boldsymbol{Q}^{-1}$ contains at most $\left\lfloor \frac{t}{2} \right\rfloor$ nonsynchronous error bursts of length up to m.

For simplicity, we also introduce the following notation for various families of matrices $\boldsymbol{Q}$ and vectors $\boldsymbol{e}$.

The family of matrices $\boldsymbol{Q}$:

- We say that a matrix $\boldsymbol{Q}$ belongs to family $\mathcal{Q}_1$ if $\boldsymbol{Q} = \operatorname{diag}(\boldsymbol{M})$ is a binary matrix of size $mn \times mn$, where $\operatorname{diag}(\boldsymbol{M})$ means a block-diagonal $mn \times mn$ matrix the main diagonal of which contains nonsingular lower triangular matrices $\boldsymbol{M}_i$ of sizes $m_i \times m_i$, $m+1 \le m_i \le 2m+2$, $\sum m_i = mn$.
- We say that a matrix $\boldsymbol{Q}$ belongs to family $\mathcal{Q}_2$ if $\boldsymbol{Q} = \operatorname{diag}(\boldsymbol{M})$ is a binary matrix of size $mn \times mn$, where $\operatorname{diag}(\boldsymbol{M})$ means a block-diagonal $mn \times mn$ matrix the main diagonal of which contains nonsingular matrices $\boldsymbol{M}_i$, and for any two adjacent matrices $\boldsymbol{M}_{i_1}$ and $\boldsymbol{M}_{i_2}$ on the main diagonal

$$\begin{pmatrix} \boldsymbol{M}_{i_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}_{i_2} \end{pmatrix}$$

that have sizes $m_{i_1} \times m_{i_1}$ and $m_{i_2} \times m_{i_2}$, the following holds:
  - $m_{i_1} + m_{i_2} = 2m$;
  - In the matrix $\boldsymbol{Q}$, matrices of sizes $m_{i_1} \times m_{i_1}$ and $m_{i_2} \times m_{i_2}$ alternate;
  - Let $\boldsymbol{M}_1$ be of size $m_1 \times m_1$ and $\boldsymbol{M}_2$ of size $m_2 \times m_2$; then, if $m_1 < m_2$, in each block of two consecutive matrices

$$\begin{pmatrix} \boldsymbol{M}_{i_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}_{i_2} \end{pmatrix}$$

the matrix of a larger size is upper triangular. If $m_1 > m_2$, then the matrices of a larger size is lower triangular.

Note that the main difference between matrices $\boldsymbol{Q}$ from the families $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is that matrices from $\mathcal{Q}_1$ obey constraints on the structure of blocks $\boldsymbol{M}_i$ (they must be lower triangular), while the choice of sizes m_i of each of the blocks remains quite flexible: $m+1 \le m_i \le 2m+2$, $\sum m_i = mn$. Elements of $\mathcal{Q}_2$ obey constraints on both the sizes of adjacent matrices $m_{i_1} + m_{i_2} = 2m$ and the structure of larger matrices $\boldsymbol{M}_i$. Below it will be shown that the constraints imposed on the structure of matrices from $\mathcal{Q}_2$ make it possible to add errors of larger weights at the encryption stage than in the case where the encryption uses matrices from $\mathcal{Q}_1$.

The family of vectors $\boldsymbol{e}$:

- We say that a vector $\boldsymbol{e}$ belongs to family $\mathcal{V}_1$ if $\boldsymbol{e}$ contains up to $\left\lfloor \frac{t}{4} \right\rfloor$ nonsynchronous error bursts of length up to m;
- We say that a vector $\boldsymbol{e}$ belongs to family $\mathcal{V}_2$ if $\boldsymbol{e}$ contains up to $\left\lfloor \frac{t}{3} \right\rfloor$ nonsynchronous error bursts of length up to m;
- We say that a vector $\boldsymbol{e}$ belongs to family $\mathcal{V}_3$ if $\boldsymbol{e}$ contains up to $\left\lfloor \frac{t}{2} \right\rfloor$ nonsynchronous error bursts of length up to m.

Consistency of $\boldsymbol{e}$ and $\boldsymbol{Q} \in \mathcal{Q}_1$. A key stage in designing a cryptosystem presented in Section 5.2 is choosing of a matrix $\boldsymbol{Q}$ which is a component of the public and secret keys and of an error vector $\boldsymbol{e}$ introduced at the encryption stage. In order to make the introduced families of matrices $\mathcal{Q}_1$ and $\mathcal{Q}_2$ consistent with the types of error vectors $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}_3$, we prove several lemmas.

Lemma 1. *Let a vector $\boldsymbol{e}$ be a nonsynchronous m-burst. Let $\boldsymbol{e}' = \boldsymbol{e}\boldsymbol{Q}^{-1}$, where $\boldsymbol{Q} \in \mathcal{Q}_1$. Then $\boldsymbol{e}'$ contains at most four synchronous m-bursts.*

Proof. Let us consider the worst case. Define a binary vector $\boldsymbol{e}$ of length mn such that this vector contains $mn - m$ zeros and the error burst has a starting coordinate which is a multiple of $m-1$. Let, for simplicity, this packet consist of ones. Then, if we represent $\boldsymbol{e}$ as $\boldsymbol{e} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$, $\boldsymbol{e}_i = (e_{i_1}, \ldots, e_{i_m})$, $e_{i_j} \in \mathbb{F}_2$, the vector $\boldsymbol{e}$ contains two consecutive vectors $\boldsymbol{e}_i$ and $\boldsymbol{e}_{i+1}$ such that $\boldsymbol{e}_i = (0, 0, \ldots, 0, 1)$, $\boldsymbol{e}_{i+1} = (1, 1, \ldots, 1, 0)$, and $\mathrm{wt}(\boldsymbol{e}_{i+1}) = m-1$. All the other $\boldsymbol{e}_j$, $j \notin \{i, i+1\}$, are zero vectors of length m. Let the matrix $\boldsymbol{Q}^{-1}$ corresponding to the nonzero segment of the vector $\boldsymbol{e}$ contain two matrices $\boldsymbol{M}_{i_1}$ and$\boldsymbol{M}_{i_2}$ with sizes $m_{i_1} \times m_{i_1}$ and $m_{i_2} \times m_{i_2}$, respectively. Consider the vectors $\boldsymbol{e}'_i = (0, 0, \ldots 0, \boldsymbol{e}_i)$ and $\boldsymbol{e}'_{i+1} = (\boldsymbol{e}_{i+1}, 0, \ldots, 0)$ of lengths m_{i_1} and m_{i_2}, respectively. When calculating $\boldsymbol{e}' = \boldsymbol{e}\boldsymbol{Q}^{-1}$, the segment of $\boldsymbol{e}'$ corresponding to the product of $(\boldsymbol{e}'_i, \boldsymbol{e}'_{i+1})$ and $\boldsymbol{Q}^{-1}$ has the form

$$(\widehat{\boldsymbol{e}}_i, \widehat{\boldsymbol{e}}_{i+1}) = (\boldsymbol{e}'_i\boldsymbol{M}_{i_1}, \boldsymbol{e}'_{i+1}\boldsymbol{M}_{i_2}).$$

Since the vector $\widehat{\boldsymbol{e}}_i$ contains one in the last position, $\boldsymbol{e}'_i\boldsymbol{M}_{i_1}$ coincides with the last row of the matrix $\boldsymbol{M}_{i_1}$, which is of weight at most m_{i_1}. In the worst case (from the point of view of error propagation), the vector $\widehat{\boldsymbol{e}}_i$ starts with 1. The vector $\widehat{\boldsymbol{e}}_{i+1}$ is a product of a vector containing $m-1$ ones in the beginning, all the other $m_{i_2} - m + 1$ symbols being 0. Thus, $\boldsymbol{e}'_{i+1}\boldsymbol{M}_{i_2}$ is of the form

$$\widehat{\boldsymbol{e}}_{i+1} = (\widehat{e}_{i+1,1}, \widehat{e}_{i+1,2}, \ldots, \widehat{e}_{i+1,m-1}, 0, \ldots, 0),$$

where $\widehat{e}_{i+1,m-1}$ can be nonzero. In the worst case, $\widehat{e}_{i+1,m-1} = 1$.

Thus, the vector $(\widehat{\boldsymbol{e}}_i, \widehat{\boldsymbol{e}}_{i+1})$ of length $m_{i_1} + m_{i_2}$, where $2m+2 \le m_{i_1} + m_{i_2} \le 4m+4$, contains an error burst of length up to $m_{i_1} + m - 1 \le 3m+1$. Obviously, this error burst is covered by at most four synchronous m-bursts, which means that the transformation $\boldsymbol{Q}^{-1}$ makes the weight (in the q-ary Hamming metric) of the error vector, which must then be decoded by the code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$, at most 4 times as large.

If in the matrix $\boldsymbol{Q}^{-1}$ the segment of the vector $\boldsymbol{e}$ where $(\boldsymbol{e}_i, \boldsymbol{e}_{i+1})$ are located corresponds to a unique matrix $\boldsymbol{M}_i$ of size at most $2m+2$, then the vector $(\boldsymbol{e}'_i, \boldsymbol{e}'_{i+1})\boldsymbol{M}_i$ covers no more than four symbols of the field $\mathbb{F}_q$.

By the construction of the matrix $\boldsymbol{Q}$, no such error burst of length at most m in the segment $(\boldsymbol{e}_i, \boldsymbol{e}_{i+1})$ of length $2m$ of the vector $\boldsymbol{e}$ can have more than two corresponding block submatrices $\boldsymbol{M}_{i_1}$ and $\boldsymbol{M}_{i_2}$.

Thus, no such error burst of length m can cover after the transformation $\boldsymbol{Q}^{-1}$ more than four symbols of $\mathbb{F}_q$. $\triangle$

Thus, if $\boldsymbol{Q} \in \mathcal{Q}_1$ and $\boldsymbol{e} \in \mathcal{V}_1$, then $\mathcal{A}(\boldsymbol{Q}, \boldsymbol{e})$ holds.

Now let us show what upper constraints must be imposed on the m_i in matrices $\boldsymbol{Q} \in \mathcal{Q}_1$ so that an arbitrary error burst of length up to m covers after the transformation $\boldsymbol{Q}^{-1}$ as few symbols of the field $\mathbb{F}_q$ as possible, which would increase the number of added errors. The lower constraint $m_i \geq m+1$ remains valid to ensure that no error burst of length at most m in the segment $(\boldsymbol{e}_i, \boldsymbol{e}_{i+1})$ of length $2m$ can correspond to more than two block matrices $\boldsymbol{M}_{i_1}$ and $\boldsymbol{M}_{i_2}$ in $\boldsymbol{Q}^{-1}$. Recall that in the worst case, multiplication of the error burst by the matrix $\boldsymbol{Q}^{-1}$ generates a burst of length $m_i + m - 1 \geq 2m$. It is clear that such a burst can cover no more than three consecutive symbols of $\mathbb{F}_q$. For an error burst length of at most $2m+1$, the number of consecutive covered field symbols for vectors of length m in $\boldsymbol{e}\boldsymbol{Q}^{-1}$ is not greater than three. Thus, if we obtain an upper constraint on m_i from

$$m_i + m - 1 \leq 2m + 1,$$

i.e., $m_i \leq m + 2$, then instead of adding $\left\lfloor \frac{t}{4} \right\rfloor$ error bursts of length up to m to the vector $\boldsymbol{e}$ it is possible to add $\left\lfloor \frac{t}{3} \right\rfloor$ error bursts of length up to m. Note that adding $\left\lfloor \frac{t}{2} \right\rfloor$ errors does not guarantee that the vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ is decodable for the above-described structure of $\boldsymbol{Q}$.

Thus, we have the following.

Lemma 2. *Let a vector $\boldsymbol{e}$ be a nonsynchronous m-burst. Let $\boldsymbol{e}' = \boldsymbol{e}\boldsymbol{Q}^{-1}$, where $\boldsymbol{Q} \in \mathcal{Q}_1$, and let the sizes m_i of the blocks $\boldsymbol{M}_i$ satisfy the inequality $m + 1 \leq m_i \leq m + 2$. Then $\boldsymbol{e}'$ contains no more than three synchronous m-bursts.*

Thus, if $m + 1 \leq m_i \leq m + 2$, $\boldsymbol{Q} \in \mathcal{Q}_1$, and $\boldsymbol{e} \in \mathcal{V}_2$, then $\mathcal{A}(\boldsymbol{Q}, \boldsymbol{e})$.

In the general case, when adding no more than $\left\lfloor \frac{t}{\ell} \right\rfloor$ error bursts, $\ell \geq 2$, of length up to m, for the vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ to be decodable (i.e., to ensure the consistency of the vector $\boldsymbol{e}$ and the matrix $\boldsymbol{Q} \in \mathcal{Q}_1$) the lower constraint $m_i \geq m + 1$ implies the upper constraint

$$m_i \leq (\ell - 1)m - m + 2.$$

Consistency of $\boldsymbol{e}$ and $\boldsymbol{Q} \in \mathcal{Q}_2$. It was shown above that sizes m_i of blocks $\boldsymbol{M}_i$ of matrices $\boldsymbol{Q} \in \mathcal{Q}_1$ significantly affect the number of bursts of errors that can be added during encryption. It is also clear that in the absence of constraints on the indices of starting positions of error bursts, up to $\left\lfloor \frac{t}{2} \right\rfloor$ error bursts of length up to m can be corrected, where t is the number of errors corrected by the $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ code. However, it was shown above that under the only constraint $m_i \geq m + 1$, where the m_i are sizes of square matrices entering the matrix $\boldsymbol{Q} \in \mathcal{Q}_1$, the largest number of error bursts added at the encryption stage cannot be greater than $\left\lfloor \frac{t}{3} \right\rfloor$. Only in this case it is possible to guarantee their correction by the code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$ after applying the transformation $\boldsymbol{Q}^{-1}$.

Let us show that if $\boldsymbol{Q} \in \mathcal{Q}_2$ and, moreover, for the sizes m_{i_1} and m_{i_2} of any two adjacent nonsingular matrices $\boldsymbol{M}_{i_1}$ and $\boldsymbol{M}_{i_2}$ we have $m_{i_1} + m_{i_2} = 2m$, then taking into account the constraints on large matrices in the $\mathcal{Q}_2$ family, the matrix $\boldsymbol{Q}$ is consistent with $\boldsymbol{e} \in \mathcal{V}_3$; i.e., at the encryption stage it would be possible to add the maximum number, $\left\lfloor \frac{t}{2} \right\rfloor$, of nonsynchronous error bursts.

Lemma 3. *If $\boldsymbol{Q} \in \mathcal{Q}_2$ and a vector $\boldsymbol{e}$ is a nonsynchronous m-burst, then the vector $\boldsymbol{e}\boldsymbol{Q}^{-1}$ contains no more than two synchronous m-bursts.*

Proof. It is clear that for $\boldsymbol{e}\boldsymbol{Q}^{-1}$ to contain no more than two synchronous m-bursts, it is necessary and sufficient that multiplying the vector $(\boldsymbol{e}_i, \boldsymbol{e}_{i+1})$ of length $2m$, which contains an error burst of length up to m at arbitrary m consecutive positions, by the corresponding section of length $2m$ of the matrix $\boldsymbol{Q}^{-1}$ does not lead to "propagation" of bursts.

Clearly, this is achieved in the case where the corresponding section of $\boldsymbol{Q}^{-1}$ is of the form

$$\begin{pmatrix} \boldsymbol{M}_{i_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}_{i_2} \end{pmatrix},$$

where $\boldsymbol{M}_{i_1}$ and $\boldsymbol{M}_{i_2}$ are square matrices of sizes $m_{i_1} \times m_{i_1}$ and $m_{i_2} \times m_{i_2}$, and $m_{i_1} + m_{i_2} = 2m$. If at the same time the matrices of sizes $m_{i_1} \times m_{i_1}$ and $m_{i_2} \times m_{i_2}$ alternate in $\boldsymbol{Q}$, then no error burst of weight m will be at the intersection of more than two matrices in $\boldsymbol{Q}$.

If, moreover, the additional constraints on the structure of larger matrices $\boldsymbol{M}_i$ are fulfilled (they are upper triangular if the size of the first block matrix $\boldsymbol{M}_1$ is less than the size of $\boldsymbol{M}_2$, and they are lower triangular if the size of the first block matrix $\boldsymbol{M}_1$ is greater than the size of $\boldsymbol{M}_2$), then, whatever the burst lying in $(\boldsymbol{e}_i, \boldsymbol{e}_{i+1})$, being multiplied by $\boldsymbol{Q}^{-1}$ it does not "propagate" to adjacent symbols, and therefore the vector $\boldsymbol{e}' = \boldsymbol{e}\boldsymbol{Q}^{-1}$ will have the same structure as the vector $\boldsymbol{e}$ generated at the encryption stage.

The only difference between the vectors $\boldsymbol{e}'$ and $\boldsymbol{e}$ is that the lengths of error bursts in $\boldsymbol{e}'$ can amount to $2m$, but the Hamming weight of $\boldsymbol{e}'$ calculated over the field $\mathbb{F}_q$ will not exceed t, which guarantees its decodability by the code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$. $\triangle$

Consistency of $\boldsymbol{e}$ and $\boldsymbol{Q}$: the main result. Combining Lemmas 1–3, we formulate a theorem connecting the families $\mathcal{Q}_1$ and $\mathcal{Q}_2$ with the families $\mathcal{V}_1$, $\mathcal{V}_2$, and $\mathcal{V}_3$ of vectors $\boldsymbol{e}$ so that $\boldsymbol{e}\boldsymbol{Q}^{-1}$ does not contain more than t synchronous error bursts of length up to m, i.e., is decodable by the code $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$.

Theorem 1. *The following statements are valid*:

- *If* $\boldsymbol{Q} \in \mathcal{Q}_1$, $\boldsymbol{e} \in \mathcal{V}_1$, *and for all blocks* $\boldsymbol{M}_i$ *of sizes* $m_i \times m_i$ *we have* $m + 1 \le m_i \le 2m + 2$ *and* $\sum m_i = mn$, *then* $\mathcal{A}(\boldsymbol{Q}, \boldsymbol{e})$ *holds*;
- *If* $\boldsymbol{Q} \in \mathcal{Q}_1$, $\boldsymbol{e} \in \mathcal{V}_1 \cup \mathcal{V}_2$, *and for all blocks* $\boldsymbol{M}_i$ *of sizes* $m_i \times m_i$ *we have* $m + 1 \le m_i \le m + 2$ *and* $\sum m_i = mn$, *then* $\mathcal{A}(\boldsymbol{Q}, \boldsymbol{e})$ *holds*;
- *If* $\boldsymbol{Q} \in \mathcal{Q}_2$ *and* $\boldsymbol{e} \in \mathcal{V}_1 \cup \mathcal{V}_2 \cup \mathcal{V}_3$, *then* $\mathcal{A}(\boldsymbol{Q}, \boldsymbol{e})$ *holds.*

Thus, at the stage of the designing of a cryptographic system based on the binary image of a generalized Reed–Solomon code, presented in Section 5.2, the developer chooses a corresponding pair $(\boldsymbol{Q}, \boldsymbol{e})$ in accordance with Theorem 1. The choice of a pair allows to control the flexibility of parameters that specify the matrix $\boldsymbol{Q}$ and the number of errors added at the encryption stage.

It should be specially noted that in contrast to the classical McEliece cryptosystem, where the public generator matrix defines a linear code equivalent to the secret one, this is not the case in our situation: right multiplication of the generator matrix $\boldsymbol{G}'_b$ by $\boldsymbol{Q} \in \mathcal{Q}_1 \cup \mathcal{Q}_2$ defines a transformation of columns in $\boldsymbol{G}'_b$, so $\boldsymbol{S}\boldsymbol{G}'_b\boldsymbol{Q}$ is not a generator matrix of the binary image of $GRS_{n,k}(\boldsymbol{\alpha}, \boldsymbol{v})$. The equivalence of codes would be preserved if $\boldsymbol{Q}$ were a block permutation of length n, the block length being m, i.e., if $\boldsymbol{Q}$ defined a permutation of symbols of $\mathbb{F}_q$. Since the code C_b corrects the maximum number of error bursts of length up to m, this set of bursts will be not correctable by the code with generator matrix $\boldsymbol{S}\boldsymbol{G}'_b\boldsymbol{Q}$ with high probability, which is a key factor underlying the proposed cryptosystem.

Next, we consider decoding attacks on the proposed class of cryptosystems.

5.4. Analysis of Some Attacks

When considering attacks, we depart from the fact that up to $\left\lfloor \frac{t}{4} \right\rfloor$ error bursts of length up to m are added at the encryption stage to the vector $\boldsymbol{e}$, although all the obtained results can easily be generalized for an arbitrary number of bursts $\left\lfloor \frac{t}{\ell} \right\rfloor$, $\ell \ge 2$.

Direct attacks. Recall that direct attacks are reduced to the exhaustive search over either information vectors $\boldsymbol{x}$ or error vectors $\boldsymbol{e}$. The complexity of direct attacks can be estimated from above as follows:

- The maximum number of rounds to recover $\boldsymbol{x}$: $\min\{2^{mk}, 2^{m(n-k)}\}$, in each round a vector of length mk or $m(n-k)$ is multiplied by a public parity check or generator matrix, which requires $m^2k(n-k)$ operations;
- The maximum number of rounds to recover $\boldsymbol{e}$: $\binom{mn}{\lfloor t/4 \rfloor}2^{m-1}$, in each round a vector $\boldsymbol{e}$ is subtracted from a received vector $\boldsymbol{y}$ and the syndrome is calculated, which requires $m^2k(n-k)$ operations.

Thus, the complexity C_{dir} of a direct attack can be estimated as

$$C_{\text{dir}} = \mathcal{O}\left(m^2k(n-k)\cdot\min\left\{2^{mk},\ 2^{m(n-k)},\ \binom{mn}{\lfloor \frac{t}{4} \rfloor}2^{m-1}\right\}\right).$$

Attacks based on information set decoding. It is known that for the classical McEliece cryptosystem, the information set decoding attack is the most effective; it determines the complexity of breaking the cryptosystem and influences the choice of code parameters (and hence the length of the public and secret keys) required for achieving a given security level.

Since $\left\lfloor \frac{t}{4} \right\rfloor$ error bursts of length up to m are added to the vector $\boldsymbol{e}$ of length mn, to find an error-free information set one has to find $\left\lfloor \frac{t}{4} \right\rfloor$ starting positions of each of the error bursts and assume that the length of each error burst is equal to m. The number of rounds to find these positions is not greater than $\binom{mn}{\lfloor t/4 \rfloor}$. Thus, the complexity of finding an error-free information set is

$$C_{\text{ISD}} = \binom{mn}{\lfloor \frac{t}{4} \rfloor}m^2k(n-k).$$

If we look for an information set among the complement of the set of $\left\lfloor \frac{t}{4} \right\rfloor$ disjoint error bursts of length m, then the last estimate can be refined:

$$C_{\text{ISD}} = \frac{m(n-1)(m(n-1)-m)(m(n-1)-2m)\dots\left(m(n-1)-m\left\lfloor \frac{t}{4} \right\rfloor\right)}{\left(\left\lfloor \frac{t}{4} \right\rfloor\right)!}m^2k(n-k).$$

Syndrome attack. The essence of the syndrome attack is to calculate the public parity check matrix $\boldsymbol{H}$ from the public generator matrix $\boldsymbol{G}$ and then reduce the problem of finding $\boldsymbol{x}$ from the expression $\boldsymbol{y} = \boldsymbol{x}\boldsymbol{G} + \boldsymbol{e}$ to solving the corresponding syndrome equation by multiplying both parts by $\boldsymbol{H}^{\mathrm{T}}$. Since

$$\boldsymbol{G} = \boldsymbol{S}\boldsymbol{G}'_b\boldsymbol{Q},$$

we have

$$\boldsymbol{H} = \boldsymbol{L}\boldsymbol{H}'_b(\boldsymbol{Q}^{-1})^{\mathrm{T}},$$

where $\boldsymbol{H}'_b$ is a parity check matrix corresponding to the generator matrix $\boldsymbol{G}'_b$ and $\boldsymbol{L}$ is some nonsingular matrix of size $m(n-k)\times m(n-k)$ over $\mathbb{F}_2$. It is clear that $\boldsymbol{L}$ does not affect the code properties. Therefore, we will assume that $\boldsymbol{L} = \boldsymbol{I}$. In this case the syndrome $\boldsymbol{Z}$ of a ciphertext $\boldsymbol{y}$ has the form

$$\boldsymbol{Z} = \boldsymbol{y}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{e}\boldsymbol{Q}^{-1}(\boldsymbol{H}'_b)^{\mathrm{T}}.$$

However, due to the randomness in the choice of $\boldsymbol{Q}$, the parity check matrix $\boldsymbol{H}$ can correspond to a code with distance much smaller than the distance of the code C_b. Thus, using a syndrome attack does not guarantee finding the vector of error bursts generated at the encryption stage.

5.5. Key Lengths

We will upper estimate the complexity C_{comp} of cryptanalysis of the cryptosystem proposed in this paper by the quantity

$$C_{\text{comp}} = \mathcal{O}(\min\{C_{\text{dir}}, C_{\text{ISD}}\}).$$

Thus, to obtain a given security level W of the system it is necessary to choose an (n,k) code (probably, truncated) $GRS_{n,k}(\boldsymbol{\alpha},\boldsymbol{v})$ over $\mathbb{F}_q$ so that to ensure $W \le C_{\text{comp}}$. Then the length of the public key will be $L_{\text{pub}} = knm^2$. Thus, an optimal (in terms of the key length) cryptosystem having security level W is determined by a triple (n,k,m) of parameters, $n \le 2^m - 1 = q - 1$, $k < n$, for which

$$\begin{cases} knm^2 \to \min, \\ C_{\text{comp}} \ge W, \\ n \le 2^m - 1, \\ 0 < k < n. \end{cases}$$

We consider several examples.

Example 1. Let $W = 2^{72}$. Consider the truncated generalized $(76, 18)$ Reed–Solomon code over the field $\mathbb{F}_q$, $q = 2^7$, obtained from the generalized Reed–Solomon code over $\mathbb{F}_q$. This code has distance 59 and corrects any 29 independent q-ary errors. Further, consider the binary image of this code by representing each element of $\mathbb{F}_q$ as a binary vector of length 7. This results in a binary $(532, 126)$ code C correcting up to 29 error bursts of length up to 7. If we take this code as a basis for constructing the cryptosystem described above, then the complexity of cryptanalysis of the system is estimated as follows:

1. $2^{mk}m^2k(n-k) > 2^{141}$ is the complexity of the brute force attack on information vectors;
2. $2^{m(n-k)}m^2k(n-k) > 2^{421}$ is the complexity of the brute force attack on information vectors for the dual code;
3. $\binom{mn}{\lfloor t/4 \rfloor}2^{m-1}m^2k(n-k) > 2^{72}$ is the complexity of the brute force attack on all error vectors;
4. The complexity of the information set decoding attack is estimated as

$$C_{\text{ISD}} = \frac{7 \cdot 75 \cdot (7 \cdot 75 - 7) \cdot (7 \cdot 75 - 14) \cdot \ldots \cdot (7 \cdot 75 - 49)}{7!} \cdot 49 \cdot 18 \cdot 58 > 2^{75}.$$

Thus, the security level of the cryptosystem is $W_c \approx 2^{72} \approx W$. In this case, the key length is $L_{\text{pub}} = 76 \cdot 18 \cdot 7^2 = 67032$, which is more than 8 times smaller than the key length of the McEliece cryptosystem based on the $(1024, 524, 101)$ Goppa code and having security level 2^{72}.

The "completeness" of the error set according to equation (5) is estimated from below as

$$\tau_{GRS} = \frac{\log_2 |\mathcal{E}|}{m(n-k)} = \frac{\log_2\left(\binom{mn}{\lfloor t/4 \rfloor}2^{m-1}\right)}{m(n-k)} \approx 0.1405,$$

which is significantly inferior to the estimate of this value for the McEliece cryptosystem $\tau_{\text{ME}} \approx 0.5681$. First of all, this indicates that the binary image of the generalized Reed–Solomon code is capable of correcting a much wider error set than the set generated in this cryptosystem. This means that it might be possible to further reduce the length of the public key, which will actually be done in the examples below.

Let us give one more example of cryptosystem parameters under the assumption that at the encryption stage $\left\lfloor \frac{t}{3} \right\rfloor$ error bursts of length up to m are added. Recall that the sizes of the matrices $\boldsymbol{M}_i$ are chosen from the set $\{m+1, m+2\}$. In this case, estimates for the complexity of cryptanalysis are obviously obtained from similar relations in the case of adding $\left\lfloor \frac{t}{4} \right\rfloor$ bursts.

Example 2. Let $W = 2^{72}$. Consider the generalized $(63, 15)$ Reed–Solomon code over the field $\mathbb{F}_q$, $q = 2^6$, obtained from the generalized Reed–Solomon code over $\mathbb{F}_q$. This code has distance 48 and corrects any 24 independent q-ary errors. Further, consider the binary image of this code obtained by representing each element of $\mathbb{F}_q$ as a binary vector of length 6. This results in a binary $(378, 90)$ code C correcting up to 24 error bursts of length up to 6. If we take this code as a basis for constructing the cryptosystem described above, then the complexity of the cryptanalysis of the system is estimated as follows:

1. $2^{mk}m^2k(n-k) > 2^{104}$ is the complexity of the brute force attack on information vectors;
2. $2^{m(n-k)}m^2k(n-k) > 2^{302}$ is the complexity of the brute force attack on information vectors for the dual code;
3. $\binom{mn}{\lfloor t/3 \rfloor}2^{m-1}m^2k(n-k) > 2^{72}$ is the complexity of the brute force attack on all error vectors;
4. The complexity of the information set decoding attack is estimated as

$$C_{\mathrm{ISD}} = \frac{6 \cdot 62 \cdot (6 \cdot 62 - 6) \cdot (6 \cdot 62 - 12) \cdot \ldots \cdot (6 \cdot 62 - 96)}{8!} \cdot 36 \cdot 15 \cdot 48 > 2^{75}.$$

Thus, the security level of the cryptosystem is $W_c \approx 2^{72} \approx W$. In this case, the key length is $L_{\mathrm{pub}} = 63 \cdot 15 \cdot 6^2 = 34020$, which is more than 15 times smaller than the key length of the McEliece cryptosystem based on the $(1024, 524, 101)$ Goppa code and having security 2^{72}.

The "completeness" of the error set according to equation (5) is estimated from below as

$$\tau_{GRS} = \frac{\log_2 |\mathcal{E}|}{m(n-k)} = \frac{\log_2\left(\binom{mn}{\lfloor t/3 \rfloor}2^{m-1}\right)}{m(n-k)} \approx 0.19147,$$

which is still less than that of the McEliece cryptosystem but greater than that for the cryptosystem based on the truncated generalized $(76, 18)$ Reed–Solomon code.

To conclude this section, we consider another example of cryptographic parameters assuming that $\left\lfloor \frac{t}{2} \right\rfloor$ error bursts of length up to m are added at the encryption stage. Recall that in this case there are some additional constraints on the matrix $\boldsymbol{Q}$, which were considered above. Then estimates of the complexity of cryptanalysis are obviously obtained from similar relations for the case of adding $\left\lfloor \frac{t}{4} \right\rfloor$ or $\left\lfloor \frac{t}{3} \right\rfloor$ error bursts.

Example 3. Let $W = 2^{72}$. Consider the truncated generalized $(46, 10)$ Reed–Solomon code over the field $\mathbb{F}_q$, $q = 2^6$, obtained from the generalized Reed–Solomon code over $\mathbb{F}_q$. This code has distance 37 and corrects any 18 independent q-ary errors. Further, consider the binary image of this code obtained by representing each element of $\mathbb{F}_q$ as a binary vector of length 6. This results in a binary $(276, 60)$ code C correcting up to 18 error bursts of length up to 6. If we take this code as a basis for constructing the cryptosystem described above, then the complexity of cryptanalysis of the system is estimated as follows:

1. $2^{mk}m^2k(n-k) > 2^{73}$ is the complexity of the brute force attack on information vectors;
2. $2^{m(n-k)}m^2k(n-k) > 2^{229}$ is the complexity of the brute force attack on information vectors for the dual code;
3. $\binom{mn}{\lfloor t/2 \rfloor}2^{m-1}m^2k(n-k) \approx 2^{73}$ is the complexity of the brute force attack on all error vectors;
4. The complexity of the information set decoding attack is estimated as

$$C_{\mathrm{ISD}} = \frac{6 \cdot 45 \cdot (6 \cdot 45 - 6) \cdot (6 \cdot 45 - 12) \cdot \ldots \cdot (6 \cdot 45 - 63)}{9!} \cdot 36 \cdot 10 \cdot 36 > 2^{74}.$$

Thus, the security level of the cryptosystem is $W_c \approx 2^{73} > W$. In this case, the key length is $L_{\text{pub}} = 46 \cdot 10 \cdot 6^2 = 16560$, which is more than 32 times smaller than the key length of the McEliece cryptosystem based on the $(1024, 524, 101)$ Goppa code and having security 2^{72}.

The "completeness" of the error set according to equation (5) is estimated from below as

$$\tau_{GRS} = \frac{\log_2 |\mathcal{E}|}{m(n-k)} = \frac{\log_2 \left(\binom{mn}{\lfloor t/2 \rfloor} 2^{m-1} \right)}{m(n-k)} \approx 0.2746,$$

which is still less than that of the McEliece cryptosystem but greater than that for the cryptosystem based on the truncated $(76, 18)$ and $(63, 15)$ generalized Reed–Solomon codes.

6. CONCLUSION

In this paper, we have considered the general problem of designing a public key cryptosystem based on error-correcting codes.

We have formulated conditions which a coding cryptosystem must obey to guarantee a required security level.

We have proposed a criteria τ to compare cryptosystems, which evaluates the interrelation between the cardinality of the error set added to the cryptosystem to provide its security and the number of check symbols of the code underlying the cryptosystem. Thus, the quantity $1 - \tau$ may be considered as the measure of a potential to improve the cryptosystem by further increasing the error set.

To demonstrate the theoretical possibility of the construction of cryptosystems for which meeting the conditions formulated in the paper is possible, we have described a scheme obeying the proposed conditions. This construction is based on binary images of generalized Reed–Solomon codes. We have shown that the construction has a smaller key length for given security parameters as compared to the McEliece cryptosystem based on binary Goppa codes.

As a result of analysis of the cryptosystem properties, we have shown that the most promising are cryptosystems where the number of added errors is much greater than half the minimum distance, while information set decoding is no more the most efficient decoding attack strategy.

The results of the paper show that designing coding cryptosystems based on the use of masking vectors (error vectors) of small Hamming weight is not efficient, since such systems are sensitive to attacks based on searching for an error-free information set. The use of error vectors that are not the lightest in their cosets can significantly reduce the efficiency of the information set decoding attack. Thus, the problem arises of finding transformations that map light representatives of cosets of codes into heavier ones. To the best of our knowledge, this problem has not been solved in coding theory so far. We hope that this problem may turn out to be useful both in cryptography and in other applications of error-correcting coding theory.

FUNDING

The paper uses results of the project "Development of Methods for Reliable and Holistic Information Transmission in Multiple Access Systems with Forward Error Correction and Digital Watermarks" carried out within the framework of the HSE University Basic Research Program in 2021.

The research of V.R. Sidorenko was supported by the European Research Council under the Horizon 2020 Program for Research and Innovation, grant no. 801434.

REFERENCES

1. McEliece, R.J., A Public-Key Cryptosystem Based on Algebraic Coding Theory, *JPL DSN Progress Report, Jet Propulsion Lab., California Inst. of Technology, Pasadena, CA*, 1978, no. 42-44, pp. 114–116. Available at https://tmo.jpl.nasa.gov/progress_report2/42-44/44N.PDF
2. Kabatianskii, G., Krouk, E., and Smeets, B., A Digital Signature Scheme Based on Random Error-Correcting Codes, *Crytography and Coding (Proc. 6th IMA Int. Conf. on Cryptography and Coding, Cirencester, UK, Dec. 17–19, 1997)*, Darnell, M., Ed., Lect. Notes Comput. Sci., vol. 1355, Berlin: Springer, 1997, pp. 161–167. https://doi.org/10.1007/BFb0024461
3. Rivest, R.L., Shamir, A., and Adleman, L.M., A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, *Commun. ACM*, 1978, vol. 21, no. 2, pp. 120–126. https://doi.org/10.1145/359340.359342
4. El Gamal, T., A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms, *IEEE Trans. Inform. Theory*, 1985, vol. 31, no. 4, pp. 469–472. https://doi.org/10.1109/TIT.1985.1057074
5. Véron, P., Code Based Cryptography and Steganography, *Algebraic Informatics (Proc. 5th Int. Conf. on Algebraic Informatics (CAI'2013), Porquerolles, France, Sept. 3–6, 2013)*, Muntean, T., Poulakis, D., and Rolland, R., Eds., Lect. Notes Comput. Sci., vol. 8080, Berlin: Springer, 2013, pp. 9–46. https://doi.org/10.1007/978-3-642-40663-8_5
6. Berger, T.P., Cayrel, P.L., Gaborit, P., and Otmani, A., Reducing Key Length of the McEliece Cryptosystem, *Progress in Cryptology — AFRICACRYPT 2009 (Proc. 2nd Int. Conf. on Cryptology in Africa, Gammarth, Tunisia, June 21–25, 2009)*, Preneel, B., Ed., Lect. Notes Comput. Sci., vol. 5580. Berlin: Springer, 2009, pp. 77–97. https://doi.org/10.1007/978-3-642-02384-2_6
7. Faugère, J.-C., Otmani, A., Perret, L., and Tillich, J.-P., Algebraic Cryptanalysis of McEliece Variants with Compact Keys, *Advances in Cryptology — EUROCRYPT 2010 (Proc. 29th Annu. Int. Conf. on the Theory and Applications of Cryptographic Techniques, French Riviera, May 30–June 3, 2010)*, Gilbert, H., Ed., Lect. Notes Comput. Sci., vol. 6110, Berlin: Springer, 2010, pp. 279–298. https://doi.org/10.1007/978-3-642-13190-5_14
8. Kocher, P.C., Timing Attacks on Implementations of Diffie–Hellman, RSA, DSS, and Other Systems, *Advances in Cryptology — CRYPTO'96 (Proc. 16th Annu. Int. Cryptology Conf., Santa Barbara, CA, USA, Aug. 18–22, 1996)*, Koblitz, N., Ed., Lect. Notes Comput. Sci., vol. 1109, Berlin: Springer, 1996, pp. 104–113. https://doi.org/10.1007/3-540-68697-5_9
9. Barker, E., NIST Special Publication (SP) 800-57 Part 1 Revision 4, *Recommendation for Key Management* – Part 1: *General*, National Inst. of Standards and Technology, Gaithersburg, MD, USA, 2016. https://doi.org/10.6028/NIST.SP.800-57pt1r4
10. Chen, L., Jordan, S., Liu, Y.-K., Moody, D., Peralta, R., Perlner, R., and Smith-Tone, D., *Report on Post-Quantum Cryptography*, NIST Internal Report 8105, National Inst. of Standards and Technology, Gaithersburg, MD, USA, 2016. https://doi.org/10.6028/NIST.IR.8105
11. Shor, P.W., Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, *SIAM Rev.*, 1999, vol. 41, no. 2, pp. 303–332. https://doi.org/10.1137/S0036144598347011
12. Berlekamp, E., McEliece, R., and van Tilborg, H., On the Inherent Intractability of Certain Coding Problems, *IEEE Trans. Inform. Theory*, 1978, vol. 24, no. 3, pp. 384–386. https://doi.org/10.1109/TIT.1978.1055873
13. Eisenbarth, T., Kumar, S., Paar, C., Poschmann, A., and Uhsadel, L., A Survey of Lightweight-Cryptography Implementations, *IEEE Des. Test Comput.*, 2007, vol. 24, no. 6, pp. 522–533. https://doi.org/10.1109/MDT.2007.178
14. Ivanov, F., Krouk, E., and Kreshchuk, A., On the Lightweight McEliece Cryptosystem for Low-Power Devices, in *Proc. 2019 XVI Int. Symp. "Problems of Redundancy in Information and Control Systems" (REDUNDANCY), Moscow, Russia, Oct. 21–25, 2019*, pp. 133–138. https://doi.org/10.1109/REDUNDANCY48165.2019.9003324

15. Misoczki, R., Tillich, J.-P., Sendrier, N., and Barreto, P.S.L.M., MDPC-McEliece: New McEliece Variants from Moderate Density Parity-Check Codes, in *Proc. 2013 IEEE Int. Symp. on Information Theory (ISIT'2013), Istanbul, Turkey, July 7–12, 2013*, pp. 2069–2073. https://doi.org/10.1109/ISIT.2013.6620590

16. Baldi, M., Chiaraluce, F., Garello, R., and Mininni, F., Quasi-cyclic Low-Density Parity-Check Codes in the McEliece Cryptosystem, in *Proc. 2007 IEEE Int. Conf. on Communications (ICC'2007), Glasgow, UK, June 24–28, 2007*, pp. 951–956. https://doi.org/10.1109/ICC.2007.161

17. Kruk, E.A., Decoding Complexity Bound for Linear Block Codes *Probl. Peredachi Inf.*, 1989, vol. 25, no. 3, pp. 103–107 [*Probl. Inf. Transm.* (Engl. Transl.), 1989, vol. 25, no. 3, pp. 251–254]. http://mi.mathnet.ru/eng/ppi665

18. Sidelnikov, V.M. and Shestakov, S.O., On Insecurity of Cryptosystems Based on Generalized Reed–Solomon Codes, *Diskret. Mat.*, 1992, vol. 4, no. 3, pp. 57–63 [*Discrete Math. Appl.* (Engl. Transl.), 1992, vol. 2, no. 4, pp. 439–444]. https://doi.org/10.1515/dma.1992.2.4.439

19. Bernstein, D.J., Lange, T., and Peters, C., Attacking and Defending the McEliece Cryptosystem, *Post-Quantum Cryptography (Proc. 2nd Int. Workshop on Post-Quantum Cryptography [PQCrypto 2008], Cincinnati, OH, USA, Oct. 17–19, 2008)*, Buchmann, J. and Ding, J., Eds., Lect. Notes Comput. Sci., vol. 5299, Berlin: Springer, 2008, pp. 31–46. https://doi.org/10.1007/978-3-540-88403-3_3

20. Becker, A., Joux, A., May, A., and Meurer, A., Decoding Random Binary Linear Codes in $2^{n/20}$: How $1 + 1 = 0$ Improves Information Set Decoding, *Advances in Cryptology — EUROCRYPT 2012 (Proc. 31st Annu. Int. Conf. on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, Apr. 15–19, 2012)*, Pointcheval, D. and Johansson, T., Eds., Lect. Notes Comput. Sci., vol. 7237, Berlin: Springer, 2012, pp. 520–536. https://doi.org/10.1007/978-3-642-29011-4_31