Problems of Information Transmission, Vol. 39, No. 2, 2003, pp. 159–169. Translated from Problemy Peredachi Informatsii, No. 2, 2003, pp. 3–14.

INFORMATION THEORY AND CODING THEORY

Codes in the Vandermonde $\mathcal{F}$-Metric and Their Application[1]

E. M. Gabidulin and V. A. Obernikhin

Moscow Institute of Physics and Technology (State University), Institutskii per. 9, Dolgoprudnyi, 141700 Russia

Received March 19, 2002; in final form, February 19, 2003

Abstract—$\mathcal{F}$-metrics are metrics based on projective sets. In this paper, construction of optimal codes for a special $\mathcal{F}$-metric associated with a generalized Vandermonde matrix is given. Encoding and fast decoding algorithms are described. A public-key cryptosystem is considered as an example of a possible application of codes constructed.

1. INTRODUCTION

In algebraic coding theory, numerous works are devoted to codes in the Hamming metric, as well as in the rank metric. Other metrics, for instance, the $\mathcal{F}$-metrics suggested in [2–5], are not examined closely; see, e.g., [6, 7], etc. Nevertheless, these metrics can provide new possibilities for both correcting special types of errors and applications in other fields, for example, in cryptography.

In this paper, we consider a class of $\mathcal{F}$-metrics associated with generalized Vandermonde matrices. For this class, it is possible to develop an interesting theory.

The paper is organized as follows. General properties of $\mathcal{F}$-metrics are described in Section 2 (based on [2]). An $\mathcal{F}$-metric associated with a generalized Vandermonde matrix is introduced in Section 3. Properties of codes in this metric and a fast decoding algorithm are presented. Possible application of such codes to public-key cryptography is described in Section 4. Lemmas whose proofs involve cumbersome computations are postponed to the Appendix.

2. $\mathcal{F}$-METRICS

2.1. General Properties

All the definitions and statements for $\mathcal{F}$-metrics in this section, except for Lemma 1, are borrowed from [2].

Let Ω be an n-dimensional vector space $\mathbb{F}_q^n$ over a finite field $\mathbb{F}_q = GF(q)$.

By the *span* $\langle X \rangle$ of a subset $X \subset \Omega$, we call the minimum linear subspace $F_X \subseteq \Omega$ containing X. Let $\mathcal{F} := \{F_1, F_2, \ldots, F_N\}$ be any family of subsets $F_i \subset \Omega$ such that $\left\langle \bigcup_{i=1}^{N} F_q \right\rangle = \Omega$.

Definition 1. The *$\mathcal{F}$-norm* ($\mathcal{F}$-weight), $\mathcal{N}_{\mathcal{F}}$, of a vector $\boldsymbol{x} \in \Omega$ is the cardinality of the smallest subset I of the set $\{1, 2, \ldots, N\}$ such that $\boldsymbol{x}$ belongs to $\left\langle \bigcup_{i \in I} F_i \right\rangle$.

Definition 2. The *$\mathcal{F}$-distance* between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is the norm of their difference, i.e., $d_{\mathcal{F}}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{N}_{\mathcal{F}}(\boldsymbol{x} - \boldsymbol{y})$.

[1] Some results of the article were presented in [1].

0032-9460/03/3902-0159 $25.00

The $\mathcal{F}$-norm has the following obvious properties:

1. $\mathcal{N}_{\mathcal{F}}(\boldsymbol{x}) = 0 \Longleftrightarrow \boldsymbol{x} = \boldsymbol{0}$;
2. $\mathcal{N}_{\mathcal{F}}(\alpha\boldsymbol{x}) = \mathcal{N}_{\mathcal{F}}(\boldsymbol{x}),\ \forall \alpha \in \mathbb{F}_q \backslash \{0\},\ \forall \boldsymbol{x} \in \Omega$;
3. $\mathcal{N}_{\mathcal{F}}(\boldsymbol{x}+\boldsymbol{y}) \le \mathcal{N}_{\mathcal{F}}(\boldsymbol{x}) + \mathcal{N}_{\mathcal{F}}(\boldsymbol{y}),\ \forall \boldsymbol{x}, \boldsymbol{y} \in \Omega$.

Thus, the $\mathcal{F}$-norm is a proper metric on the space Ω. Since $\mathcal{F}$-norms with respect to families $\{F_1, F_2, \ldots, F_N\}$ and $\{\langle F_1\rangle, \langle F_2\rangle, \ldots, \langle F_N\rangle\}$ coincide,

$$\left\langle \bigcup_{i\in I} F_i \right\rangle = \left\langle \bigcup_{i\in I} \langle F_i\rangle \right\rangle,$$

we may confine our consideration to families of linear subspaces only. Moreover, if $F_i \subset F_j$ for distinct i and j, then elimination of F_i from the family does not change the norm.

Example 1. Let $N = n$, $\Omega = \mathbb{F}_q^n$, and $\mathcal{F} := \{\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_n\}$, where $\boldsymbol{E}_i$ is a standard basis in $\mathbb{F}_q^n$. Then the $\mathcal{F}$-norm is the Hamming norm: $\mathcal{N}_{\mathcal{F}}(\boldsymbol{x}) = d_H(\boldsymbol{x}),\ \forall \boldsymbol{x} \in \mathbb{F}_q^n$.

If $\mathcal{F} := \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n\}$, where the vectors $\boldsymbol{f}_i$ form a basis in $\mathbb{F}_q^n$, then this metric is equivalent to the Hamming metric.

Example 2. Let $\Omega = \mathbb{F}_q^{m\times\ell}$ be the linear space of $m \times \ell$ matrices over $\mathbb{F}_q$. Let $\mathcal{R}$ denote the set of rank-1 matrices:

$$\mathcal{R} = \left\{ \boldsymbol{M} :\ \operatorname{rank} \boldsymbol{M} = 1,\ \boldsymbol{M} \in \mathbb{F}_q^{m\times\ell} \right\}.$$

Since any matrix of rank r can be represented as a sum of r matrices of rank 1 and any sum of rank-1 matrices is a matrix of rank not greater than r, we have $\mathcal{N}_{\mathcal{R}}(\boldsymbol{A}) = \operatorname{rank} \boldsymbol{A}$, $\forall \boldsymbol{A} \in \mathbb{F}_r^{m\times\ell}$. Therefore, we get the rank metric [8].

Definition 3. Any subset $C \in \Omega$ is called a *code*.

Definition 4. The *$\mathcal{F}$-distance* of a code $C \subset \Omega$ is the integer

$$d_{\mathcal{F}}(C) := \min\{d_{\mathcal{F}}(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{x}, \boldsymbol{y} \in C,\ \boldsymbol{x} \ne \boldsymbol{y}\}.$$

Definition 5. If all elements of a family $\mathcal{F} := \{F_1, F_2, \ldots, F_N\}$ are vectors, then the metric generated by the family is called a *projective $\mathcal{F}$-metric*. In this case, we will denote elements of the family by $\boldsymbol{f}_i$, i.e., $\mathcal{F} := \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N\}$.

In the sequel, speaking about $\mathcal{F}$-metrics, we consider projective $\mathcal{F}$-metrics only.

Lemma 1 (generalized Singleton bound). *For any linear code $C \subseteq \mathbb{F}_q^n$ of dimension k, we have the following inequality:*

$$d_{\mathcal{F}}(C) \le n - k + 1. \tag{1}$$

Proof. Let $\mathcal{F} := \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N\}$. Choose a basis $\boldsymbol{f}_{i_1}, \boldsymbol{f}_{i_2}, \ldots, \boldsymbol{f}_{i_n}$ in $\mathbb{F}_q^n$ composed of vectors of the family $\mathcal{F}$. Consider the metric $\mathcal{F}_{\text{red}}$ generated by these vectors only, $\mathcal{F}_{\text{red}} := \{\boldsymbol{f}_{i_1}, \boldsymbol{f}_{i_2}, \ldots, \boldsymbol{f}_{i_n}\}$. Since this metric is equivalent to the Hamming metric, we have the usual Singleton bound: $d_{\mathcal{F}_{\text{red}}}(C) \le n - k + 1$. Adding extra vectors to $\mathcal{F}_{\text{red}}$ can only decrease the $\mathcal{F}$-norm; i.e., $d_{\mathcal{F}}(C) \le d_{\mathcal{F}_{\text{red}}}(C)$. $\triangle$

We refer to a code meeting this bound as a *code with the maximum $\mathcal{F}$-distance*.

2.2. Parent Code

Define the mapping $\varphi\colon \mathbb{F}_q^N \to \mathbb{F}_q^n$ as $\varphi(\boldsymbol{e}_i) := \boldsymbol{f}_i$, $i = 1, \ldots, N$, where $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ is a standard basis in $\mathbb{F}_q^N$ and $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N\}$ are vectors that define the $\mathcal{F}$-metric.

Definition 6. The *parent code* is the kernel $P := \ker(\varphi) \subset \mathbb{F}_q^N$.

Since the condition $\boldsymbol{x} \in \ker(\varphi)$ can be expressed as $\boldsymbol{F}\boldsymbol{x} = \boldsymbol{0}$, where $\boldsymbol{F}$ is the matrix whose columns are coordinates of the vectors $\varphi(\boldsymbol{e}_i) = \boldsymbol{f}_i$, $i = 1, \ldots, N$, in the space $\mathbb{F}_q^n$, the parent code P is an $[N, N-n]$ code with parity-check matrix $\boldsymbol{F}$.

Let $w(D)$ be the *weight of a coset* $D \in \mathbb{F}_q^N/P$, i.e., the Hamming weight of the coset leader. Information about the $\mathcal{F}$-weight distribution of $\mathbb{F}_q^n$ is of great importance in coding theory. The following lemma allows one to find the Hamming weight distribution and reduce the calculation of the $\mathcal{F}$-norm to calculation of the Hamming spectrum of cosets of the parent code.

Lemma 2. *The $\mathcal{F}$-norm of any vector $\boldsymbol{y} \in \mathbb{F}_q^n$ is equal to the weight of a coset that has $\boldsymbol{y}$ as a syndrome:*

$$d_{\mathcal{F}}(\boldsymbol{y}) = d_H(\varphi^{-1}(\boldsymbol{y})).$$

It follows from the lemma that the maximum $\mathcal{F}$-norm is equal to the covering radius of the parent code P:

$$\rho(P) := \max\left\{w(D),\ D \in \mathbb{F}_q^N/P\right\}.$$

3. CODES IN THE VANDERMONDE $\mathcal{F}$-METRIC

3.1. Vandermonde $\mathcal{F}$-Metric

Let us define the Vandermonde $\mathcal{F}$-metric in the following way. As vectors $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N$ that define the $\mathcal{F}$-metric, take columns of a generalized Vandermonde matrix

$$\boldsymbol{F} = \begin{pmatrix} u_1 & u_2 & \ldots & u_N \\ u_1x_1 & u_2x_2 & \ldots & u_Nx_N \\ u_1x_1^2 & u_2x_2^2 & \ldots & u_Nx_N^2 \\ \ldots & \ldots & \ldots & \ldots \\ u_1x_1^{n-1} & u_1x_2^{n-1} & \ldots & u_Nx_N^{n-1} \end{pmatrix}, \tag{2}$$

where $n \le N$, $x_i \in \mathbb{F}_q = GF(q)$ are pairwise distinct, and $u_i \in \mathbb{F}_q = GF(q)$ are nonzero, $i = 1, \ldots, N$.

The parent code for this $\mathcal{F}$-metric is a generalized Reed–Solomon code (GRS code).

3.2. Codes

Our goal now is to construct a linear code with the maximum $\mathcal{F}$-distance $d_{\mathcal{F}} = n - k + 1$. Let a linear $[n, k]$ code C be defined by its transposed generator matrix

$$\boldsymbol{G}^T = \begin{pmatrix} g_{11} & g_{21} & \ldots & g_{k1} \\ g_{12} & g_{22} & \ldots & g_{k2} \\ \ldots & \ldots & \ldots & \ldots \\ g_{1n} & g_{2n} & \ldots & g_{kn} \end{pmatrix}.$$

For a set of k information symbols $\boldsymbol{a} = (a_1, a_2, \ldots, a_k)^T$, the corresponding code vector is calculated as $\boldsymbol{g} = \boldsymbol{G}^T\boldsymbol{a}$.

Let $\boldsymbol{G}^T$ be as follows:

$$\boldsymbol{G}^T = \begin{pmatrix} v_1 & v_2 & \ldots & v_k \\ v_1y_1 & v_2y_2 & \ldots & v_ky_k \\ v_1y_1^2 & v_2y_2^2 & \ldots & v_ky_k^2 \\ \ldots & \ldots & \ldots & \ldots \\ v_1y_1^{n-1} & v_2y_2^{n-1} & \ldots & v_ky_k^{n-1} \end{pmatrix}, \tag{3}$$

where $v_i \in \mathbb{F}_q$ are nonzero and $y_i \in \mathbb{F}_q = GF(q)$ are pairwise distinct. Moreover, let us choose y_i to be distinct from any of the x_i. In this case, the concatenation of the matrices $\boldsymbol{F}$ and $\boldsymbol{G}^T$ is also a generalized Vandermonde matrix. The code dimension k must satisfy the inequality $k + N \le q$ since the maximum possible number of columns of a generalized Vandermonde matrix over $GF(q)$ equals q.

Lemma 3. *The code C defined by the matrix $\boldsymbol{G}^T$ is a code with the maximum $\mathcal{F}$-distance: $d_{\mathcal{F}}(C) = n - k + 1$. Consequently, the code corrects up to $t_k = \left[\dfrac{n-k}{2}\right]$ $\mathcal{F}$-errors.*

Proof. Let $\boldsymbol{g}$ be a code vector obtained as the product of the matrix $\boldsymbol{G}^T$ and an arbitrary vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_k)^T$ of Hamming weight $s \neq 0$, i.e., $\boldsymbol{g}^T = \boldsymbol{a}^T\boldsymbol{G}$. Then $\boldsymbol{g}$ can be represented as a linear combination of columns of $\boldsymbol{G}^T$ corresponding to the nonzero components of $\boldsymbol{a}$: $\boldsymbol{g} = a_{j_1}\boldsymbol{g}_{j_1} + a_{j_2}\boldsymbol{g}_{j_2} + \ldots + a_{j_s}\boldsymbol{g}_{j_s}$. Let ℓ denote the $\mathcal{F}$-weight of $\boldsymbol{g}$. By the definition of the $\mathcal{F}$-norm, $\boldsymbol{g}$ can be represented as $\boldsymbol{g} = b_1\boldsymbol{f}_{i_1} + b_2\boldsymbol{f}_{i_2} + \ldots + b_\ell\boldsymbol{f}_{i_\ell}$, where none of the b_i equals zero. It follows from the equation $\boldsymbol{g} = b_1\boldsymbol{f}_{i_1} + b_2\boldsymbol{f}_{i_2} + \ldots + b_\ell\boldsymbol{f}_{i_\ell} = a_{j_1}\boldsymbol{g}_{j_1} + a_{j_2}\boldsymbol{g}_{j_2} + \ldots + a_{j_s}\boldsymbol{g}_{j_s}$ that $\ell + s$ *distinct* columns $\boldsymbol{f}_{i_1}, \boldsymbol{f}_{i_2}, \ldots, \boldsymbol{f}_{i_\ell}, \boldsymbol{g}_{j_1}, \boldsymbol{g}_{j_2}, \ldots, \boldsymbol{g}_{j_s}$ of the generalized Vandermonde matrix are linearly dependent. Therefore, $\ell + s \ge n + 1$, or

$$\mathcal{N}_{\mathcal{F}}(\boldsymbol{g}) \ge n - s + 1. \tag{4}$$

Thus, for the minimum $\mathcal{F}$-distance of the code, we have $d_{\mathcal{F}}(C) \ge n - k + 1$. Taking into account the generalized Singleton bound (1), we conclude that $d_{\mathcal{F}}(C) = n - k + 1$. $\triangle$

3.3. Fast Decoding

We reduce decoding in the $\mathcal{F}$-metric to decoding GRS codes. In turn, for GRS codes, fast decoding algorithms exist.

Let $\boldsymbol{c} = \boldsymbol{g} + \boldsymbol{e}$, where $\boldsymbol{g}$ is a code vector and $\boldsymbol{e}$ is an error. Let t denote the $\mathcal{F}$-weight of the error. Then $\boldsymbol{e}$ can be represented as a linear combination of the vectors $\{\boldsymbol{f}_i\}$

$$\boldsymbol{e} = m_1\boldsymbol{f}_1 + m_2\boldsymbol{f}_2 + \ldots + m_N\boldsymbol{f}_N$$

such that $d_H(\boldsymbol{m}) = t$, where $\boldsymbol{m} = (m_1, m_2, \ldots, m_N)^T$.

Let us show that there exists a fast decoding algorithm if $t \le t_k$. Consider the concatenation of the matrices $\boldsymbol{F}$ and $\boldsymbol{G}^T$:

$$\left(\boldsymbol{F} \mid \boldsymbol{G}^T\right) = \left(\begin{array}{cccc|cccc} u_1 & u_2 & \ldots & u_N & v_1 & v_2 & \ldots & v_k \\ u_1x_1 & u_2x_2 & \ldots & u_Nx_N & v_1y_1 & v_2y_2 & \ldots & v_ky_k \\ u_1x_1^2 & u_2x_2^2 & \ldots & u_Nx_N^2 & v_1y_1^2 & v_2y_2^2 & \ldots & v_ky_k^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ u_1x_1^{n-1} & u_2x_2^{n-1} & \ldots & u_Nx_N^{n-1} & v_1y_1^{n-1} & v_2y_2^{n-1} & \ldots & v_ky_k^{n-1} \end{array}\right)$$

$$= \left(\begin{array}{cccc|cccc} u_1 & u_2 & \ldots & u_N & u_{N+1} & u_{N+2} & \ldots & u_{N+k} \\ u_1x_1 & u_2x_2 & \ldots & u_Nx_N & u_{N+1}x_{N+1} & u_{N+2}x_{N+2} & \ldots & u_{N+k}x_{N+k} \\ u_1x_1^2 & u_2x_2^2 & \ldots & u_Nx_N^2 & u_{N+1}x_{N+1}^2 & u_{N+2}x_{N+2}^2 & \ldots & u_{N+k}x_{N+k}^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ u_1x_1^{n-1} & u_2x_2^{n-1} & \ldots & u_Nx_N^{n-1} & u_{N+1}x_{N+1}^{n-1} & u_{N+2}x_{N+2}^{n-1} & \ldots & u_{N+k}x_{N+k}^{n-1} \end{array}\right), \tag{5}$$

where we used the notations $x_{N+i} = y_i$ and $u_{N+i} = v_i$, $i = 1, 2, \ldots, k$.

Let $\boldsymbol{R}$ denote a nonsingular square matrix of order n formed by the last n columns of matrix (5). Let us premultiply (5) by $\boldsymbol{R}^{-1}$ and thus reduce it to the canonical form:

$$\boldsymbol{R}^{-1}\left(\boldsymbol{F} \mid \boldsymbol{G}^T\right) = \left(\widetilde{\boldsymbol{F}} \mid \widetilde{\boldsymbol{G}}^T\right) = \left(\begin{array}{cc|c} \boldsymbol{B}_1 & \boldsymbol{E}_{n-k} & 0 \\ \boldsymbol{B}_2 & 0 & \boldsymbol{E}_k \end{array}\right), \tag{6}$$

where $\boldsymbol{E}_\ell$ is the identity matrix of order ℓ and the $n \times (N-n+k)$ matrix $\left(\begin{smallmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}_2\end{smallmatrix}\right)$ is a generalized Cauchy matrix, whose elements are of the form $b_{ij} = \dfrac{\alpha_i \beta_j}{\mu_i - \nu_j}$ and can be obtained explicitly (see Lemma A.1).

Let us premultiply $\boldsymbol{c} = \boldsymbol{g} + \boldsymbol{e}$ by $\boldsymbol{R}^{-1}$:

$$\boldsymbol{R}^{-1}(\boldsymbol{g}+\boldsymbol{e}) = \boldsymbol{R}^{-1}(\boldsymbol{g}+\boldsymbol{F}\boldsymbol{m}) = \widetilde{\boldsymbol{g}} + \widetilde{\boldsymbol{F}}\boldsymbol{m} = \widetilde{\boldsymbol{g}} + \widetilde{\boldsymbol{e}}.$$

The first $n-k$ components of the vector $\widetilde{\boldsymbol{g}} = \boldsymbol{R}^{-1}\boldsymbol{g}$ are zero:

$$\widetilde{\boldsymbol{g}} = (0, 0, \ldots, 0, \widetilde{g}_{n-k+1}, \widetilde{g}_{n-k+2}, \ldots, \widetilde{g}_n)^T.$$

This allows us to find the first $n-k$ components of the vector $\widetilde{\boldsymbol{e}} = \boldsymbol{R}^{-1}\boldsymbol{F}\boldsymbol{m} = \widetilde{\boldsymbol{F}}\boldsymbol{m}$. Let us show that, given these components, it is possible to reconstruct the vector $\boldsymbol{m}$. To this end, we have to solve the system of equations $\widetilde{\boldsymbol{F}}\boldsymbol{m} = \widetilde{\boldsymbol{e}}$:

$$\begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{E}_{n-k} \\ \boldsymbol{B}_2 & 0 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{pmatrix} = \begin{pmatrix} \widetilde{e}_1 \\ \widetilde{e}_2 \\ \vdots \\ \widetilde{e}_{n-k} \\ * \\ \vdots \\ * \end{pmatrix}. \tag{7}$$

Consider the first $n-k$ rows of system (7):

$$\begin{pmatrix} b_{1,1} & b_{1,2} & \ldots & b_{1,N+k-n} & 1 & \ldots & 0 \\ b_{2,1} & b_{2,2} & \ldots & b_{2,N+k-n} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ b_{n-k,1} & b_{n-k,2} & \ldots & b_{n-k,N+k-n} & 0 & \ldots & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{pmatrix} = \begin{pmatrix} \widetilde{e}_1 \\ \widetilde{e}_2 \\ \vdots \\ \widetilde{e}_{n-k} \end{pmatrix}. \tag{8}$$

The matrix $\boldsymbol{H} = (\boldsymbol{B}_1 \mid \boldsymbol{E}_{n-k})$ is the concatenation of the generalized Cauchy matrix and the identity matrix. The matrix $\boldsymbol{H}$ can be transformed into a generalized Vandermonde matrix $\boldsymbol{H}' = (\boldsymbol{\Psi}\boldsymbol{B}_1 \mid \boldsymbol{\Psi})$ if we premultiply it by an appropriate nonsingular square matrix $\boldsymbol{\Psi}$ of order $n-k$ (see Lemma A.2). Thus, it is necessary to solve the system of equations

$$\boldsymbol{H}'\boldsymbol{m} = \boldsymbol{\Psi} \begin{pmatrix} \widetilde{e}_1 \\ \widetilde{e}_2 \\ \vdots \\ \widetilde{e}_{n-k} \end{pmatrix}, \tag{9}$$

where the right-hand side and the matrix $\boldsymbol{H}'$ are known.

Solving the system is the decoding problem for the GRS code $C_{H'}$ with the standard-form parity-check matrix $\boldsymbol{H}'$. The problem has a unique solution if the Hamming weight of the vector $\boldsymbol{m}$ does not exceed the error-correcting capability of the code, $\left[\dfrac{n-k}{2}\right]$. In this case, a fast decoding algorithms exists (see, e.g. [9]); on applying it, we find the vector $\boldsymbol{m}$ and then the vectors $\boldsymbol{e}$ and $\boldsymbol{g}$.

4. APPLICATION TO CRYPTOGRAPHY

4.1. Public-Key Cryptosystem Based on the Niederreiter System

A public-key cryptosystem that uses an error vector as a plaintext was first introduced by Niederreiter [10]. A parity-check matrix $\boldsymbol{H}$ of a GRS code premultiplied by a nonsingular matrix $\boldsymbol{S}$, which hides the structure of $\boldsymbol{H}$, is chosen as a public key:

$$\boldsymbol{H}_{\text{pub}} = \boldsymbol{S}\boldsymbol{H}.$$

A plaintext is an error vector $\boldsymbol{e}$ of Hamming weight $d_H(\boldsymbol{e}) \le \left[\dfrac{d-1}{2}\right]$, where d is the minimum distance of the GRS code. Encryption is calculating the syndrome $s = \boldsymbol{S}\boldsymbol{H}\boldsymbol{e}$. A legitimate user premultiplies the syndrome by $\boldsymbol{S}^{-1}$, applies a fast decoding algorithm for the GRS code, and finds the vector $\boldsymbol{e}$. To get the plaintext, an intruder would have to solve the problem of decoding a code with known parity-check matrix $\boldsymbol{S}\boldsymbol{H}$ in the general case. Up to now, no algorithm for solving the problem in a polynomial time is known.

However, Sidelnikov and Shestakov showed [11] that it is possible to obtain a matrix of the form $\boldsymbol{S}'\boldsymbol{H}'$ from $\boldsymbol{H}_{\text{pub}}$ in a polynomial time, where $\boldsymbol{S}'$ is a nonsingular square matrix and $\boldsymbol{H}'$ is a generalized Vandermonde matrix, which is also a parity-check matrix for the same GRS code. Thus, the original Niederreiter cryptosystem could not withstand the structural attack of reconstructing a private key by a known public one.

One of the possible ways to modify the cryptosystem is introducing a hiding matrix $\boldsymbol{X}$. In this case, a public key is $\boldsymbol{H}_{\text{pub}} = \boldsymbol{S}(\boldsymbol{H} + \boldsymbol{X})$. In [12], it was proposed to use hiding matrices of rank 1. Results of the present article allow one to use hiding matrices of much larger ranks.

The cryptosystem is constructed in the following way. First, a legitimate user chooses a matrix $\boldsymbol{F}$, whose columns define an $\mathcal{F}$-metric. The parent code with the parity-check matrix $\boldsymbol{F}$ must have a fast decoding algorithm in the Hamming metric. Then it is necessary to choose a transposed generator matrix $\boldsymbol{G}^T$ of a linear code C with a fast decoding algorithm in the $\mathcal{F}$-metric. In the case of the Vandermonde $\mathcal{F}$-metric, matrices $\boldsymbol{F}$ and $\boldsymbol{G}^T$ are (2) and (3) respectively.

Next, one chooses a nonsingular square matrix $\boldsymbol{S}$ of order n and a permutation matrix $\boldsymbol{P}$ of order N.

A *secret key* is the set of matrices $\{\boldsymbol{F}, \boldsymbol{G}^T, \boldsymbol{S}, \boldsymbol{P}\}$.

A *public key* is the matrix

$$\boldsymbol{H}_{\text{pub}} = \boldsymbol{S}(\boldsymbol{F} + \boldsymbol{G}^T\boldsymbol{U})\boldsymbol{P}, \tag{10}$$

where $\boldsymbol{U}$ is a random $k \times N$ matrix. Columns of the matrix $\boldsymbol{G}^T\boldsymbol{U}$ are code vectors $\boldsymbol{G}_i$ of the code C. The matrix $\boldsymbol{U}$ is not needed for decryption, it should only be made unavailable for a cryptanalyst.

A *plaintext* is an N-dimensional vector $\boldsymbol{m} = (m_1, m_2, \ldots, m_N)^T$, where $m_i \in \mathbb{F}_q$, and $d_H(\boldsymbol{m}) = t_{\min} = \min\{t_k, t_P\}$, where t_k is the error-correcting capability of the code defined by $\boldsymbol{G}_T$ in the space with the $\mathcal{F}$-metric and t_P is the error-correcting capability of the parent code. Hence, the number of possible messages is $C_n^{t_{\min}}(q-1)^{t_{\min}}$.

Encryption. A cyphertext is computed as the syndrome

$$\begin{aligned}\boldsymbol{c} = \boldsymbol{H}_{\text{pub}}\boldsymbol{m} = \boldsymbol{S}(\boldsymbol{F} + \boldsymbol{G}^T\boldsymbol{U})\boldsymbol{P}\boldsymbol{m} = \boldsymbol{S}(\boldsymbol{F} + \boldsymbol{G}^T\boldsymbol{U})\widetilde{\boldsymbol{m}} \\ = \boldsymbol{S}\left(\widetilde{m}_1(\boldsymbol{f}_1 + \boldsymbol{G}_1) + \widetilde{m}_2(\boldsymbol{f}_2 + \boldsymbol{G}_2) + \ldots + \widetilde{m}_N(\boldsymbol{f}_N + \boldsymbol{G}_N)\right) = \boldsymbol{S}\left(\boldsymbol{g} + \boldsymbol{e}\right),\end{aligned}$$

where $\widetilde{\boldsymbol{m}} = \boldsymbol{P}\boldsymbol{m}$ and $\boldsymbol{f}_i$ and $\boldsymbol{G}_i$ are columns of the matrices $\boldsymbol{F}$ and $\boldsymbol{G}^T\boldsymbol{U}$ respectively.

Decryption. A legitimate user premultiplies the received cyphertext $\boldsymbol{S}(\boldsymbol{g} + \boldsymbol{e})$ by $\boldsymbol{S}^{-1}$ and then, applying the fast decoding algorithm in the $\mathcal{F}$-metric, gets the vectors $\boldsymbol{g}$ and $\boldsymbol{e}$. Then he applies the

fast decoding algorithm for the parent code to $\boldsymbol{e}$ to calculate the vector $\widetilde{\boldsymbol{m}}$. It is worth noting that, in the case of the Vandermonde $\mathcal{F}$-metric, the fast decoding algorithm in the $\mathcal{F}$-metric immediately yields $\widetilde{\boldsymbol{m}}$. To get the plaintext $\boldsymbol{m}$, one has only to multiply $\widetilde{\boldsymbol{m}}$ by $\boldsymbol{P}^{-1}$.

If the code C consists of the zero vector only, we get the classical Niederreiter cryptosystem, broken by Sidelnikov and Shestakov [11].

It should be noted that not every matrix $\boldsymbol{U}$ can be used to construct a secure cryptosystem. For example, if $\boldsymbol{U} = \boldsymbol{V}\boldsymbol{F}$, then

$$\boldsymbol{H} = \boldsymbol{S}(\boldsymbol{E}_n + \boldsymbol{G}^T\boldsymbol{V})\boldsymbol{F}\boldsymbol{P}, \tag{11}$$

that is, we only get an additional scrambling matrix for the Niederreiter cryptosystem.

5. CONCLUSION

In this paper, we have constructed codes for a projective $\mathcal{F}$-metric associated with a generalized Vandermonde matrix, developed a fast decoding algorithm, and considered an application of the codes to cryptography. In the future, it seems to be of great interest to define $\mathcal{F}$-metrics for other codes and find fast decoding algorithms for them. Also, it makes sense to investigate the security of the public-key cryptosystem presented.

In conclusion, we would like to thank Alexey Ourivsky, as well as an anonymous reviewer, for their constructive criticism, which helped us to improve the paper.

APPENDIX

Lemma A.1. *Let $\boldsymbol{V}$ be an $n \times (m+n)$ generalized Vandermonde matrix and $\boldsymbol{R}$ be the matrix formed by the last n columns of $\boldsymbol{V}$. The product $\boldsymbol{R}^{-1}\boldsymbol{V}$ is the concatenation of a generalized Cauchy matrix and an identity matrix:*

$$\boldsymbol{V} = \begin{pmatrix} z_1 & z_2 & \dots & z_m & z_{m+1} & \dots & z_{m+n} \\ z_1x_1^2 & z_2x_2^2 & \dots & z_mx_m^2 & z_{m+1}x_{m+1}^2 & \dots & z_{m+n}x_{m+n}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ z_1x_1^{n-1} & z_2x_2^{n-1} & \dots & z_mx_m^{n-1} & z_{m+1}x_{m+1}^{n-1} & \dots & z_{m+n}x_{m+n}^{n-1} \end{pmatrix}. \tag{12}$$

Proof. Consider the set of Lagrange interpolation polynomials of degree $n-1$:

$$f_i(x) = \prod_{\substack{1\le s\le n \\ s\ne i}} \frac{(x_{m+s} - x)}{(x_{m+s} - x_{m+i})} = \sum_{s=1}^{n} f_{is}x^{s-1}, \quad i = 1, \dots, n. \tag{13}$$

Note that

$$f_i(x_j) = \sum_{s=1}^{n} f_{is}x_j^{s-1} = \delta_{i,j-m} = \begin{cases} 1, & i = j, \\ 0, & i \ne j, \end{cases} \qquad i = 1, \dots, n, \quad j = m+1, \dots, m+n. \tag{14}$$

Define the matrix $\boldsymbol{A}$ as $\boldsymbol{A}_{i\ell} = \left[\dfrac{f_{i\ell}}{z_{m+i}}\right]$, $i, \ell = 1, \dots, n$. Consider the product of $\boldsymbol{A}$ and $\boldsymbol{V}$:

$$[\boldsymbol{A}\boldsymbol{V}]_{ij} = \sum_{s=1}^{n} \frac{f_{is}z_jx_j^{s-1}}{z_{m+i}} = \frac{z_j}{z_{m+i}} \sum_{s=1}^{n} f_{is}x_j^{s-1} = \frac{z_j}{z_{m+i}} f_i(x_j).$$

For $j = m+1 \dots, m+n$, we have

$$[\boldsymbol{AV}]_{ij} = \delta_{i,j-m},$$

that is, the last n columns of $\boldsymbol{AV}$ form an identity matrix:

$$\boldsymbol{AV} = \begin{pmatrix} * & \dots & * & 1 & 0 & \dots & 0 \\ * & \dots & * & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ * & \dots & * & 0 & 0 & \dots & 1 \end{pmatrix}. \tag{15}$$

Consequently, $\boldsymbol{A} = \boldsymbol{R}^{-1}$.

For $j = 1, \dots, m$, we have

$$\begin{aligned} \left[\boldsymbol{R}^{-1}\boldsymbol{V}\right]_{ij} &= \frac{z_j f_i(x_j)}{z_{m+i}} = \frac{z_j}{z_{m+i}} \prod_{\substack{1 \le s \le n \\ s \ne i}} \frac{(x_{m+s} - x_j)}{(x_{m+s} - x_{m+i})} \\ &= \frac{z_j}{z_{m+i}} \frac{\prod\limits_{1 \le s \le n} (x_{m+s} - x_j)}{\prod\limits_{\substack{1 \le s \le n \\ s \ne i}} (x_{m+s} - x_{m+i})} \frac{1}{x_{m+i} - x_j} \\ &= \left(\frac{1}{z_{m+i} \prod\limits_{\substack{1 \le s \le n \\ s \ne i}} (x_{m+s} - x_{m+i})} \right) \left(z_j \prod_{1 \le s \le n} (x_{m+s} - x_j) \right) \frac{1}{x_{m+i} - x_j} \\ &= \frac{\alpha_i \beta_j}{x_{m+i} - x_j}, \end{aligned}$$

where

$$\begin{aligned} \alpha_i &= \frac{1}{z_{m+i} \prod\limits_{\substack{1 \le s \le n \\ s \ne i}} (x_{m+s} - x_{m+i})} \ne 0, \quad i = 1, 2, \dots, n, \\ \beta_j &= z_j \prod_{1 \le s \le n} (x_{m+s} - x_j) \ne 0, \quad j = 1, 2, \dots, m. \end{aligned} \tag{16}$$

Thus, the product $\boldsymbol{R}^{-1}\boldsymbol{V}$ is the concatenation of a generalized Cauchy matrix and an identity matrix.

Lemma A.2. *Let us be given a concatenation of a generalized Cauchy matrix and an identity matrix*

$$\boldsymbol{C} = \begin{pmatrix} \dfrac{a_1 b_1}{c_1 - d_1} & \dfrac{a_1 b_2}{c_1 - d_2} & \dots & \dfrac{a_1 b_m}{c_1 - d_m} & 1 & 0 & \dots & 0 \\ \dfrac{a_2 b_1}{c_2 - d_1} & \dfrac{a_2 b_2}{c_2 - d_2} & \dots & \dfrac{a_2 b_m}{c_2 - d_m} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dfrac{a_n b_1}{c_n - d_1} & \dfrac{a_n b_2}{c_n - d_2} & \dots & \dfrac{a_n b_m}{c_n - d_m} & 0 & 0 & \dots & 1 \end{pmatrix}$$

with given coefficients a_i, c_i, b_j, d_j, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$. Then there exists a nonsingular matrix $\boldsymbol{\Psi}$ of order n such that the product $\boldsymbol{\Psi C}$ is a generalized Vandermonde matrix.

Proof. Define the elements x_ℓ and z_ℓ, $\ell = 1, \ldots, m+n$, of the matrix $\boldsymbol{V}$ (12) in the following way:

$$\begin{aligned} x_j &= d_j, & x_{m+i} &= c_i, \\ z_j &= \frac{b_j}{\prod\limits_{1\le s\le n}(x_{m+s} - x_j)}, & z_{m+i} &= \frac{1}{a_i \prod\limits_{\substack{1\le s\le n \\ s\ne i}}(x_{m+s} - x_{m+i})}, \\ j &= 1, \ldots, m, & i &= 1, \ldots, n. \end{aligned} \tag{17}$$

Since all the x_ℓ are distinct and z_ℓ are nonzero, $\boldsymbol{V}$ is a generalized Vandermonde matrix. If $\boldsymbol{R}$ is a matrix formed by the last n columns of $\boldsymbol{V}$, then, by the previous lemma, the product $\boldsymbol{R}^{-1}\boldsymbol{V}$ is the concatenation of a generalized Cauchy matrix and an identity matrix. Taking (16) and (17) into account, we get $\boldsymbol{R}^{-1}\boldsymbol{V} = \boldsymbol{C}$. Thus, $\boldsymbol{\Psi} = \boldsymbol{R}$. $\triangle$

It is obvious that the matrix $\boldsymbol{\Psi}$ can be defined in different ways. For example, for any $\gamma \ne 0$, it is possible to define x_ℓ and z_ℓ, $\ell = 1, \ldots, m+n$, as

$$\begin{aligned} x_j &= \gamma d_j, & x_{m+i} &= \gamma c_i, \\ z_j &= \frac{b_j}{\prod\limits_{1\le s\le n}(x_{m+s} - x_j)}, & z_{m+i} &= \frac{1}{\gamma a_i \prod\limits_{\substack{1\le s\le n \\ s\ne i}}(x_{m+s} - x_{m+i})}, \\ j &= 1, \ldots, m, & i &= 1, \ldots, n. \end{aligned} \tag{18}$$

According to Lemma A.2, a generalized Vandermonde matrix $\boldsymbol{V}$ with coefficients (18) can be transformed into a concatenation of a generalized Cauchy matrix and an identity matrix. It follows from (16) that $\alpha_i = \gamma a_i$, $\beta_j = b_j$, and

$$\frac{\alpha_i\beta_j}{x_{m+i} - x_{m+j}} = \frac{\gamma a_i b_j}{\gamma c_i - \gamma d_j} = \frac{a_i b_j}{c_i - d_j}.$$

Lemma A.3. *Let us be given a matrix* $\left(\boldsymbol{F} \mid \boldsymbol{G}^T\right)$ *of the form* (5)*. The matrix* $\boldsymbol{H}'$ *used in the fast decoding algorithm (see Section* 3.3*) can be calculated according to the following formulas:*

$$\boldsymbol{H}' = \begin{pmatrix} \xi_1 x_1 & \xi_1 x_2 & \ldots & \xi_1 x_N \\ \xi_2 x_2^2 & \xi_2 x_2^2 & \ldots & \xi_2 x_N^2 \\ \ldots & \ldots & \ldots & \ldots \\ \xi_{n-k} x_1^{n-k-1} & \xi_{n-k} x_2^{n-k-1} & \ldots & \xi_{n-k} x_N^{n-k-1} \end{pmatrix},$$

where

$$\begin{aligned} \xi_j &= z_j \prod_{n-k+1\le s\le n}(x_{N-n+k+s} - x_j), \quad j = 1, \ldots, N-n+k, \\ \xi_{N-n+k+i} &= z_{N-n+k+i} \prod_{n-k+1\le s\le n}(x_{N-n+k+s} - x_{N-n+k+i}), \quad i = 1, \ldots, n-k. \end{aligned}$$

Proof. It is easily seen that the coefficients x_ℓ, $\ell = 1, \ldots, N$, for $\boldsymbol{H}'$ can be taken the same.

To obtain formulas for the coefficients ξ_ℓ, we use (16) and (17):

$$\begin{aligned}
\xi_j &= \frac{b_j}{\prod\limits_{1\le s\le n-k}(x_{N-n+k+s}-x_j)} \\
&= \frac{z_j \prod\limits_{1\le s\le n}(x_{N-n+k+s}-x_j)}{\prod\limits_{1\le s\le n-k}(x_{N-n+k+s}-x_j)} \\
&= z_j \prod_{n-k+1\le s\le n}(x_{N-n+k+s}-x_j), \quad j=1,\dots,N-n+k, \\
\xi_{N-n+k+i} &= \frac{1}{a_i \prod\limits_{\substack{1\le s\le n-k\\ s\ne i}}(x_{N-n+k+s}-x_{N-n+k+i})} \\
&= \frac{z_{N-n+k+i} \prod\limits_{\substack{1\le s\le n\\ s\ne i}}(x_{N-n+k+s}-x_{N-n+k+i})}{\prod\limits_{\substack{1\le s\le n-k\\ s\ne i}}(x_{N-n+k+s}-x_{N-n+k+i})} \\
&= z_{N-n+k+i} \prod_{n-k+1\le s\le n}(x_{N-n+k+s}-x_{N-n+k+i}), \quad i=1,\dots,n-k. \quad \triangle
\end{aligned}$$

It is obvious that the matrix $\boldsymbol{\Psi}$ (see Section 3.3) consists of the last $n-k$ columns of $\boldsymbol{H}'$.

REFERENCES

1. Gabidulin, E.M. and Obernikhin, V.A., Vandermonde and $\mathcal{F}$-Metrics, in *Proc. 8th Int. Workshop on Algebraic and Combinatorial Coding Theory, Tsarskoe Selo, Russia, 2002*, pp. 124–127.
2. Gabidulin, E.M. and Simonis, J., Metrics Generated by Families of Subspaces, *IEEE Trans. Inf. Theory*, 1998, vol. 44, no. 5, pp. 1336–1341.
3. Sharma, B.D. and Kaushik, M.L., On Algebra of Sharma and Kaushik's Metric Inducing Partitions of $\mathbb{Z}_q$, *J. Combin. Inf. Syst. Sci.*, 1986, vol. 11, pp. 1–14.
4. Gabidulin, E.M., Combinatorial Metrics in Coding Theory, in *Proc. 2nd Int. Sympos. on Information Theory, Moscow–Yerevan, 1971*, pp. 39–43.
5. Gabidulin, E.M. and Bossert, M., Codes Resistant to the Phase Rotation, in *Proc. 4th Sympos. on Communication and Applications, Charlotte Mason College, Lake District, UK, July 1997*, pp. 253–257.
6. Gabidulin, E.M. and Bossert, M., Hard and Soft Decision Decoding of Phase Rotation Invariant Block Codes, in *Proc. 1998 Int. Zurich Seminar on Broadband Communications: Accessing, Transmission, Networking. ETH Zurich, Switzerland, February, 1998.*
7. Gabidulin, E.M. and Simonis, J., Perfect Codes for Metrics Generated by Primitive 2-Error-Correcting Binary BCH Codes, *Probl. Peredachi Inf.*, 1999, vol. 35, no. 3, pp. 40–47 [*Probl. Inf. Trans.* (Engl. Transl.), 1999, vol. 35, no. 3, pp. 224–230].
8. Gabidulin, E.M., Theory of Codes with Maximal Rank Distance, *Probl. Peredachi Inf.*, 1985, vol. 21, no. 1, pp. 3–16 [*Probl. Inf. Trans.* (Engl. Transl.), 1985, vol. 21, no. 1, pp. 1–12].
9. MacWilliams, F.J. and Sloane, N.J.A., *The Theory of Error-Correcting Codes*, Amsterdam: North-Holland, 1977. Translated under the title *Teoriya kodov, ispravlyayushchikh oshibki*, Moscow: Svyaz', 1979.
10. Niederreiter, H., Knapsack-Type Cryptosystem and Algebraic Coding Theory, *Probl. Control Inf. Theory*, vol. 15, no. 2, 1986, pp. 159–166.

11. Sidelnikov, V.M. and Shestakov, S.O., On Insecurity of Cryptosystems Based on Generalized Reed–Solomon Codes, *Diskr. Mat.*, 1992, vol. 4, no. 3, pp. 57–63.

12. Gabidulin, E., Ourivski, A., and Pavlouchkov, V., On the Modified Niederreiter Cryptosystem, in *Proc. Information Theory and Networking Workshop, Metsovo, Greece, 1999*, p. 50.