Code-Based Public-Key Cryptosystem Based on Bursts-Correcting Codes

E. Krouk, A. Ovchinnikov
Saint-Petersburg State University of Aerospace Instrumentation
Saint-Petersburg, Russia
email: ekrouk@vu.spb.ru, mldoc@ieee.org

Abstract—**In this paper, the public-key cryptosystems based on error-correcting codes are considered. The most known code-based public-key cryptosystem belongs to McEliece and its security is based on decoding vectors of given weight t in linear code, equivalent to some private code with minimal distance $d = 2t + 1$. Another class of code-based cryptosystems is known, whose security is based on complete decoding task (or searching through all possible error vectors). It is supposed that the security of these systems may significantly overcome those of McEliece. In the paper, the cryptosystem from this class is proposed based on bursts-correcting codes.**

Keywords–Code-based public-key cryptosystems; Cryptosystems based on complete decoding task; Bursts-correcting codes.

I. Introduction

The public-key cryptosystem, proposed in 1978 by R. J. McEliece, is based on error-correcting codes [1]. The idea of the system is to select the error-correcting code for which the effective decoding algorithm is known, and then to hide this code in linear code of arbitrary structure. The description of initial code, usually given by its generator matrix, serves as private key, while the description of obtained code with arbitrary structure is public key. Being very computationally effective, McEliece cryptosystem did not obtain much practical usage, which is mainly argued by its large key sizes.

Though the decoding task of arbitrary linear code is NP-complete [2], it should be noted that in classical variant of McEliece cryptosystem its public and private keys are equivalent codes, and in fact the adversary should solve the task of decoding in sphere of some radius, which seems simpler than arbitrary linear code decoding by minimal distance. Besides, the attacks revealing the code's structure are also possible [3]-[4].

In [5], the class of public-key cryptosystems is proposed which is based on the task of complete decoding, that is, decoding of coset leaders in the standard array [6]. However, the selection of particular system from this class requires definition of masking transformation and the set of error vectors applied during encryption. Some attacks on the variants of such definitions were considered in [7]. Practical examples of the systems based on complete decoding task are also given in [7][8].

The paper is organized as follows. Section II gives the description of McEliece cryptosystem. Section III describes the class of cryptosystems based on complete decoding task. In Section IV, the variant of the system from this class is proposed based on bursts-correcting codes. Section V concludes the paper.

II. McEliece code-based cryptosystem

The construction of McEliece public-key cryptosystem is based on linear (n, k) code for which the polynomial decoding procedure is known, providing correction of any combination of t or less errors. The family of Goppa codes are usually considered for this purpose [1][6].

In McEliece cryptosystem, each user constructs private and correspondent public keys as follows:

1) Select integers k, n, t as general system parameters.
2) Select generator $(k \times n)$ matrix $\mathbf{G}$ of linear (n, k) code, for which the effective procedure ψ of correcting any combination of t errors is known.
3) Select random binary non-singular $(k \times k)$ matrix $\mathbf{M}$.
4) Select random $(n \times n)$ permutation matrix $\mathbf{P}$.
5) Calculate $(k \times n)$ matrix $\mathbf{G}' = \mathbf{MGP}$.
6) Public key is $(\mathbf{G}', t)$, private key is $(\mathbf{M}, \mathbf{G}, \mathbf{P})$.

To encrypt the message, one should do the following:

1) Represent the message as binary k-bit sequence $\mathbf{m}$.
2) Select random n-bit binary vector $\mathbf{e}$ of weight t.
3) Calculate ciphertext $\mathbf{c} = \mathbf{mG}' + \mathbf{e}$.

To decrypt the ciphertext, one should do the following:

1) Calculate $\mathbf{c}' = \mathbf{cP}^{-1}$.
2) Obtain $\mathbf{m}'$ by decoding $\mathbf{c}'$ in code $\mathbf{G}$ using ψ.
3) Calculate $\mathbf{m} = \mathbf{m}'\mathbf{M}^{-1}$.

Decryption is correct, since $\mathbf{c}' = \mathbf{cP}^{-1} = (\mathbf{mG}' + \mathbf{e})\mathbf{P}^{-1} = (\mathbf{mMGP} + \mathbf{e})\mathbf{P}^{-1} = (\mathbf{mM})\mathbf{G} + \mathbf{eP}^{-1}$ and $\mathbf{eP}^{-1}$ is the vector of weight t.

In the next section, we will describe the class of public-key cryptosystems based on complete decoding task.

III. The class of cryptosystems based on complete decoding task

The straightforward attack on McEliece cryptosystem is decoding of error vector of weight t in the code $\mathbf{G}'$. The system would be significantly harder break into if decryption would require to correct not only error vectors of weight t, but all coset leaders, that is, performing complete decoding [6].

Consider the following variant of public and private keys construction [5][7]:

1) Select generator $(k \times n)$ matrix $\mathbf{G}$ of linear (n, k) code, for which the effective procedure ψ of correcting any errors from the error set E is known (for example, E may be the set of vectors of weight t).
2) Select random binary non-singular $(n \times n)$ matrix $\mathbf{M}$.
3) Calculate $(k \times n)$ matrix $\mathbf{G}' = \mathbf{GM}$.
4) Define the error set $E' = \{\mathbf{e}' : \mathbf{e}' = \mathbf{eM}, \mathbf{e} \in E\}$.
5) Public key is $(\mathbf{G}', E')$, private key is $(\mathbf{G}, \mathbf{M})$.

To encrypt the message, one should do the following:

1) Represent the message as binary k-bit sequence $\mathbf{m}$.
2) Select random n-bit binary vector $\mathbf{e}' \in E'$.
3) Calculate ciphertext $\mathbf{c} = \mathbf{mG}' + \mathbf{e}'$.

To decrypt the ciphertext, one should do the following:

1) Calculate $\mathbf{c}' = \mathbf{cM}^{-1}$.

Copyright (c) IARIA, 2017.   ISBN: 978-1-61208-562-3

2) Obtain $\mathbf{m}$ by decoding $\mathbf{c}'$ in code $\mathbf{G}$ using ψ.

Decryption is correct, since $\mathbf{c}' = \mathbf{c}\mathbf{M}^{-1} = (\mathbf{m}\mathbf{G}\mathbf{M} + \mathbf{e}\mathbf{M})\mathbf{M}^{-1} = \mathbf{m}\mathbf{G} + \mathbf{e}$, where $\mathbf{e} \in E$, and the procedure ψ may be effectively used.

The security of the described class of systems is based on the fact that the adversary needs decoding in the code $\mathbf{G}'$, while $\mathbf{G}'$ is not only non-equivalent to the code $\mathbf{G}$ as in McEliece cryptosystem, but after multiplying $\mathbf{G}$ by $\mathbf{M}$ from the right the error-correction capability of public code $\mathbf{G}'$ is unknown (and may be rather low). In addition, the structure (for example, weight) of vector $\mathbf{e}'$ is unknown (and, in fact, it may not be the coset leader for $\mathbf{G}'$), thus the best attack may turn out not to perform the complete decoding, but instead to use brute force by vectors from E', which may be a more complicated task. Finally, we note that the set E itself may not be published and this may be used to further strengthen the system.

On the other hand, the described class proposes only a general approach, and not the particular cryptosystem. First, not only Goppa code may be used as private code $\mathbf{G}$, and not only vectors of fixed weight may form the set E. This is of special interest since, in the last years, code-based cryptosystems using codes other than Goppa codes have been considered [9][10]. Surely, by selecting particular classes of codes new possibilities may arise for the adversary, which should be thoroughly taken in consideration.

Next, the method of defining the set E' should be specified, which for security reasons should have exponential cardinality. In [7][8], some examples of such definition are given. In the next section, the public-key code-based cryptosystem from the described class is proposed, based on bursts-correcting codes.

IV. CODE-BASED CRYPTOSYSTEM USING BURSTS-CORRECTING CODES

Let us consider the following variant of the system from the class described in the previous section:

1) Select generator $(k \times n)$ matrix $\mathbf{G}$ of linear (n, k) code, for which the effective procedure ψ of correcting any errors from the error set E is known.
2) Select random binary non-singular $(n \times n)$ matrix $\mathbf{M}_2$.
3) Define the error set $\tilde{E}$ and $(n \times n)$ matrix $\mathbf{M}_1$ such that for any $\tilde{\mathbf{e}} \in \tilde{E}$ vector $\tilde{\mathbf{e}}\mathbf{M}_1$ belongs to E.
4) Calculate matrix $\mathbf{M} = \mathbf{M}_1\mathbf{M}_2$.
5) Calculate $(k \times n)$ matrix $\mathbf{G}' = \mathbf{G}\mathbf{M}_2$.
6) Public key is $(\mathbf{G}', \mathbf{M}, \tilde{E})$, private key is $(\mathbf{G}, \mathbf{M}_1, \mathbf{M}_2)$.

To encrypt the message, one should do the following:

1) Represent the message as binary k-bit sequence $\mathbf{m}$.
2) Select random n-bit binary vector $\tilde{\mathbf{e}} \in \tilde{E}$ and calculate $\mathbf{e}' = \tilde{\mathbf{e}}\mathbf{M}$.
3) Calculate ciphertext $\mathbf{c} = \mathbf{m}\mathbf{G}' + \mathbf{e}'$.

To decrypt the ciphertext, one should do the following:

1) Calculate $\mathbf{c}' = \mathbf{c}\mathbf{M}_2^{-1}$.
2) Obtain $\mathbf{m}$ by decoding $\mathbf{c}'$ in code $\mathbf{G}$ using ψ.

Decryption is correct, since $\mathbf{c}' = \mathbf{c}\mathbf{M}_2^{-1} = (\mathbf{m}\mathbf{G}\mathbf{M}_2 + \tilde{\mathbf{e}}\mathbf{M}_1\mathbf{M}_2)\mathbf{M}_2^{-1} = \mathbf{m}\mathbf{G} + \mathbf{e}$, where $\mathbf{e} \in E$, and the procedure ψ may be effectively used.

In this variant, the set E' is defined by the vectors $\tilde{\mathbf{e}}\mathbf{M}$, which in turn requires effective definition of $\tilde{E}$. Besides, the matrix $\mathbf{M}_1$ should be defined mapping the vectors from $\tilde{E}$ to E.

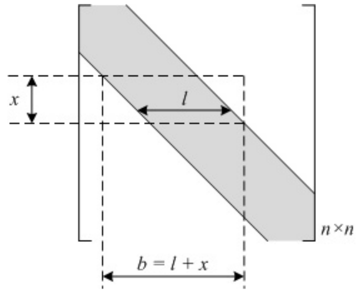

Figure 1. Definition of matrix $\mathbf{M}_1$

For example, $\tilde{E}$ and E may coincide and be formed by the vectors of some fixed weight. In this paper, we propose the system, where the set E is formed by vectors in which the number of positions between first and last non-zero elements are no greater than some value b and such error vectors are called error bursts of length b. In coding theory, the bursts-correcting codes are known to be capable of correcting single error bursts of a given length [11][12]. As the set $\tilde{E}$ we will also consider the set of error bursts of some length.

Consider as $\mathbf{M}_1$ the matrix shown in Figure 1. The gray color corresponds to positions filled by random binary elements, other positions are zeros. Clearly, such matrix defines mapping of error bursts of length x to error bursts of length b.

Then, the specification of the proposed system may be finalized by the following conditions:

- Set E: the set of bursts of length $\leq b$.
- Matrix $\mathbf{G}$ should define the code for which effective procedure of correcting single bursts of length b is known.
- Set $\tilde{E}$: the set of bursts of length $\leq x$. Clearly, x should be included in system's public key as definition of $\tilde{E}$.

Note that code $\mathbf{G}'$ has no known structure: neither its bursts-correction capability nor its minimal distance are known. Moreover, the set $E' = \{\mathbf{e}' : \mathbf{e}' = \tilde{\mathbf{e}}\mathbf{M}\}$ contains vectors of arbitrary structure, with arbitrary weights and which are not error bursts. Thus, it seems that the structure of bursts used by private code cannot be exploited by the adversary, and the best attacking strategy is to search through the elements of E'.

Quantitative estimation of system parameters: bursts lengths x and b, cardinalities of E' and $\tilde{E}$, and finally selection of k and n which define the key size depends on class of bursts-correction codes being used. This class should contain exponential number of codes for given b, k, n, and the codes should not possess the structure which may be used by the adversary to perform the structural attack. One variant for such codes is the class of low-density parity-check codes (LDPC). The bursts-correction capabilities of some LDPC codes were investigated in [13][14]. However, these codes cannot be directly applied in proposed cryptosystem since they are strongly structured, and the task of bursts-correction code selection for usage in the proposed cryptosystem may be considered as further research.

V. Conclusion

In this paper, a code-based cryptosystem using bursts-correction codes is proposed. This system belongs to the class of cryptosystems based on complete decoding task. It is supposed that cryptosystems from this class allow to achieve better security than the McEliece cryptosystem. The selection of particular codes for usage in the proposed system, which allows qualitative estimation of parameters and perhaps requires additional cryptanalysis research is the direction of further investigations.

Acknowledgment

The reported study was funded by RFBR according to the research project No. 16-01-00716 a.

References

[1] R. J. McEliece, “A Public-Key Cryptosystem Based on Algebraic Coding Theory,” 1978, DSN progress report #42-44, Jet Propulsion Laboratory, Pasadena, California.

[2] E. Berlekamp, R. McEliece, and H. van Tilborg, “On the inherent intractability of certain coding problems (corresp.),” IEEE Transactions on Information Theory, vol. 24, no. 3, May 1978, pp. 384–386.

[3] V. M. Sidelnikov and S. O. Shestakov, “On insecurity of cryptosystems based on generalized Reed-Solomon codes,” Discrete Mathematics and Applications, vol. 2, no. 4, 1992, pp. 439–444.

[4] E. Krouk, A. Ovchinnikov, and E. Vostokova, “About one modification of McEliece cryptosystem based on Plotkin construction,” in 2016 XV International Symposium Problems of Redundancy in Information and Control Systems (REDUNDANCY), Sept 2016, pp. 75–78.

[5] E. Krouk, “A New Public-Key Cryptosystem,” in Sixth Joint Swedish-Russian International Workshop on Information Theory, Moelle, Sweden, 1993, pp. 285–286.

[6] F. MacWilliams and N. Sloane, The Theory of Error-Correcting Codes. North-Holland publishing company, 1983, 782 p.

[7] E. Krouk and U. Sorger, “A Public Key Cryptosystem Based on Total Decoding of Linear Codes,” in VI International Workshop ”Algebraic and combinatorial coding theory”, Pskov, 1998, pp. 116–118.

[8] G. Kabatiansky, S. Semenov, and E. Krouk, Error Correcting Coding And Security For Data Networks: Analysis Of The Superchannel Concept. John Wiley & Sons, 2005, 278 p.

[9] R. Misoczki, J. P. Tillich, N. Sendrier, and P. S. L. M. Barreto, “MDPC-McEliece: New McEliece variants from Moderate Density Parity-Check codes,” in 2013 IEEE International Symposium on Information Theory, July 2013, pp. 2069–2073.

[10] E. M. Gabidulin, A. V. Ourivski, B. Honary, and B. Ammar, “Reducible rank codes and their applications to cryptography,” IEEE Transactions on Information Theory, vol. 49, no. 12, Dec 2003, pp. 3289–3293.

[11] R. Blahut, Theory and practice of error control codes. Addison-Wesley, 1983, 500 p.

[12] W. Zhang and J. K. Wolf, “A class of binary burst error-correcting quasi-cyclic codes,” IEEE Transactions on Information Theory, vol. 34, no. 3, May 1988, pp. 463–479.

[13] E. A. Krouk and S. V. Semenov, “Low-Density Parity-Check Burst Error-Correcting Codes,” in 2 International Workshop ”Algebraic and combinatorial coding theory”, Leningrad, 1990, pp. 121–124.

[14] E. A. Krouk and A. A. Ovchinnikov, “2-Stripes Block-Circulant LDPC Codes for Single Bursts Correction,” Smart Innovation, Systems and Technologies, vol. 55, June 2016, pp. 11–23.

[15] N. Sendrier, “Finding the permutation between equivalent linear codes: the support splitting algorithm,” IEEE Transactions on Information Theory, vol. 46, no. 4, Jul 2000, pp. 1193–1203.

[16] E. Krouk and A. Ovchinnikov, “About one structural attack on McEliece cryptosystem,” in 2016 XV International Symposium Problems of Redundancy in Information and Control Systems (REDUNDANCY), Sept 2016, pp. 71–74.