Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 26 (2023), No. 1, pp. 103–114
DOI : 10.1080/09720529.2021.1933706

New McEliece cryptosystem based on non-permutation equivalent polar codes

Belkacem Imine *
Faculty of Electrical Engineering
University of Science and Technology of Oran
Laboratory of Coding and Security of Information (LACOSI)
(USTO) 1505 Oran M'Naouer 31000
Algeria

Naima Hadj-Said †
Faculty of Computer Science
University of Science and Technology of Oran (USTO)
1505 Oran M'Naouer 31000
Algeria

Adda Ali-Pacha §
Faculty of Electrical Engineering
University of Science and Technology of Oran
Laboratory of Coding and Security of Information (LACOSI)
(USTO) 1505 Oran M'Naouer 31000
Algeria

Abstract

In this paper, we propose a new McEliece public-key cryptosystem based on polar codes with non-permutation equivalent. Due to the fact that polar code is a decreasing monomial code, any permutation of columns of its generator matrix can be found. Thus, any McEliece cryptosystem based on only the permutation equivalent of the standard polar code seems to be insecure against some structural attacks. In contrast, our proposed scheme is able

*E-mail: imine.belkacem@gmail.com, belkacem.imine@univ-usto.dz (Corresponding Author)
†E-mail: naima.hadjsaid@univ-usto.dz
§E-mail: adda.alipacha@univ-usto.dz

to resist against these attacks and achieves good security level against the generic decoding attacks.

Subject Classification: *94A60, 11T71, 14G50.*

Keywords: *Code-based cryptography, McEliece's cryptosystem, Polar code, Post-quantum cryptography, Public-key cryptosystem.*

1. Introduction

It has been shown that the emergence of quantum computers threatens the cryptosystems based on numbers theory such as RSA [22], Diffi-Helman [7], and the cryptosystems based on elliptic curve discrete logarithm problem such as elliptic curve cryptography (ECC). Later, ECC has recognized more attention especially in the Internet of Things (IoT) domain and wireless sensor network (WSN) [13]. In [28] the authors pointed out the various important aspects in the implementation of ECC. Unfortunately, all these attempts can be broken in polynomial time when quantum computers appear, which led to find out another techniques that allow to resist against any quantum computer attacks. In 1978, McEliece proposed its first cryptosystem [17], which can resist against quantum computer attacks using error-correcting codes and exactly binary Goppa codes. The security of McEliece cryptosystem based on two assumptions, the difficulty of decoding a large linear code which turned out to be an NP-complete problem [5] and there is no polynomial-time algorithm can solve this problem, and the second assumption based on the problem of distinguishing the secret code from a random code. The McEliece cryptosystem was not broken so far.

Despite these advantages, the McEliece cryptosystem is unused in practice due to the large public key size which is the generator matrix of binary Goppa code, and thus several attempts have been discovered in order to minimize the public key size.

In 1986, H. Niederreiter [20] proposed the parity-check matrix of generalized Reed-Solomon code as a public key, but [26] proved that the Niederreiter's cryptosystem is insecure. In 1994, V.M. Sidelnikov replaced the binary goppa codes by using binary Reed-Muller codes [25]. In 1996, H.Janwa and Moreno proposed a new variant of McEliece cryptosystem using algebraic-geometric codes [12]. In 2007, M. Baldi et al. succeeded to minimize the public key size using the quasi-cyclic property of the quasi-cyclic low density parity check QC-LDPC code [2]. Unfortunately, most of these attempts turned out to be insecure, see [3, 9, 18].

Two best-known attacks against any variant of the McEliece cryptosystem, the first one is a structural attack in which the adversary try to find the secret code structure from the public code, and the second one is a generic attack in which the adversary try to recover the plaintext from the cipher-text without the knowledge of the secret key. The most famous structural attack is the support splitting algorithm [23], and the most famous generic attack is Information Set Decoding, see [6, 8, 14, 15, 27].

Due to the fact that polar codes have large hulls and large permutation group such as Goppa codes, they are also proposed as new code-based cryptography scheme [10, 11, 24]. Thus an adversary cannot solve the code equivalence problem using support splitting algorithm [23]. However, Bardet et al. introduced a powerful structural attack [4], they managed to solve the code equivalence problem by searching to the low-weight codewords in the public code and determining the structure of the low-weight codewords in the private code. [16] Proposed an efficient cryptosystem using polar code in which this attack cannot be applicable by adding w column vectors to the end of the generator matrix of the private code, on the other hand, this weakens the transmission rate.

The remainder of this paper is subdivided into five sections. In section two, we introduce the original McEliece cryptosystem. In section three, we introduce the polar code and its encoding and decoding algorithms. In section four, we introduce our approach. In section five, we analyse the security level of our cryptosystem, and a conclusion is given in last section.

McEliece's cryptosystem

In [17], R. J. McEliece proposed a new public-key cryptosystem based on error correcting codes, he chose the binary Goppa codes which can correct t errors as candidate in its cryptosystem. Key generation, encryption and decryption operations are done as follows:

Keys generation

Generate the generator matrix G of the (n, k, t) Goppa code C, then generate two binary square matrix S and P of length k and n respectively, such that S is an invertible matrix and P is a permutation matrix.

- The public key $(G'=SGP\ ,t)$
- The secret key (S, G, P)

Encryption

To encrypt a secret message m of length k, the transmitter chooses a random binary error vector e of weight at most t and length n, then computes the cipher-text

$$c' = mG' + e$$

Decryption

To decipher c', the receiver decodes the vector $c'P^{-1}$ to a correct codeword c using the t-decoding algorithm of the private code C.

- $c = decA(c'P^{-1}) = decA(mSG + eP^{-1}) = mSG,$ such that eP^{-1} has weight at most t.
- Then extract the information vector $m' = mS$ from c and compute $m = m'S^{-1} = mSS^{-1}$

Polar code

Polar codes are recent optimal codes invented in 2009 by Erdal Arikan [1] that can achieve the capacity of any binary input symmetric memoryless channel (B-DMC) like the binary erasure channel BEC and binary symmetric channel BSC. The polar code's construction is based on channel polarization, where some sub-channels become noiseless which is devoted to transmit the information bits and the other become completely noisy which is devoted to transmit the frozen bits. The selection of the information bits and the frozen bits sets depends on the type of the channel $\{W : X \to Y\}$ and the transition probability $\{W(y \mid x), x \in X, y \in Y\}$, where $X = \{0,1\}$ is the input alphabet and Y is the output alphabet. Several methods are introduced to calculate the frozen bits set, among them [1, 19].

We denote $Z(W) = \Sigma_{y \in Y} \sqrt{W(y \mid 0)W(y \mid 1)}$ the Bhattacharyya parameter of the channel W which measures the channel's reliability, where $Z(W) \in [0,1]$. For BEC(ε) channel $Z(W) = \varepsilon,$ and for BSC channel $Z(W) = 2\sqrt{pq},$ where p and q are the transition probabilities of the BSC channel. We denote (N, K, A, A^c), the polar code of length N and dimension K with the frozen bit set A^c, and information bit set A. And if we consider N sub-channel's Bhattacharya parameters $\{Z(W)^i, 1 \le i \le N\}$ where i represents the sub-channel's index, then A contains the indices of the K sub-channels that have the lower Bhattacharya parameter values, and

A^c contains the indices of the *N-K* sub-channels that have the higher Bhattacharya parameter values.

If we consider the linear transform $T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{\otimes n}$ which represents the *nth* Kronoker product of the matrix $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, then the generator matrix *G* of the $[N = 2^n, K]$ polar code is the sub-matrix $G_{K\times N}$ of *T* consisting of the rows of *T* that indexed to *A*.

$$G = G_{K\times N} = \{T_{i,1..N} : i \in A\} \tag{1}$$

Encoding

To encode the binary information vector $m^K = [m_0 m_1 ... m_{K-1}]$ to $u^N = [u_0 u_1 ... u_{N-1}]$, we multiply m^K by *G*.

- $u^N = m^K G$

Decoding

In [1], Arikan published an efficient decoding algorithm of polar code entitled Successive Cancellation (SC) decoding, however [21, 29] managed to improve this algorithm.

Successive Cancellation (SC) Decoder

Using the log likelihood ratio (*LLR*), the SC decoder can obtain the transmitted information vector from the received corrupted vector r^N. The log likelihood ratio for every received bit r_i can be calculated from (equation 2).

$$L_N^{(i)}(r^N, \hat{u}^{i-1}) = \log \frac{W^{(i)}(r^N, \hat{u}^{i-1} \mid u_i = 0)}{W^{(i)}(r^N, \hat{u}^{i-1} \mid u_i = 1)} \tag{2}$$

Hence, the SC decoder can estimate $\hat{u}_i$ for every $i = 1,2....N$ as follow:

$$\hat{u}_i = \begin{cases} 0, \text{ for } i \in A \text{ and } L_N^{(i)}(r^N, \hat{u}^{i-1}) \geq 0 \\ 1, \text{ for } i \in A \text{ and } L_N^{(i)}(r^N, \hat{u}^{i-1}) < 0 \\ 0, \text{ for } i \in A^c \end{cases} \tag{3}$$

However, [29] proved that output one bit-estimation for a given received bit can be an erroneous choice, and proposed to produce two

decisions for each bit r_i not to be frozen. Then assigning for each decision a decision metric. The sum of these decision metrics builds a path metric *PM* that's represent a belief of a particular codeword. In this stage, the SCL algorithm keeps the *L* most likely codewords that have the lower path metric. In the final stage, the SCL outputs one codeword which have the lower path metric from the *L* candidate codewords.

The proposed cryptosystem

Unlike to the original McEliece's cryptosystem where the private and the public codes are permutation equivalent. In our proposal, the private and the public codes are not-permutation equivalent in which the permutation matrix *P* is replaced by an invertible sparse matrix *Q* of row and column weight *j*. However, in this case the private code must correct $t = jt'$ errors.

Keys generation

- Choose an [*N, K, t*] polar code *C* of length $N = 2^n$ and dimension *K* and which can correct *t* errors, then generate its generator matrix *G*.
- Randomly, generate an invertible scrambling matrix *S* of length *K* over the binary field $\mathbb{F}_2$.
- Randomly, generate a binary invertible sparse matrix *Q* of length *N* and row and column weight *j*, such that $j > 1,$ then generate its inverse Q^{-1} which is a dense matrix.

  The public key: $\left(G' = SGQ^{-1},\ t' = \left\lfloor \frac{t}{j} \right\rfloor\right)$

  The secret key: *S, G, Q*

Encryption

If *Bob* wants to send a message *m* to *Alice*, first he obtains the public key G' of *Alice* from the public directory, then computes $c = mG' + e,$ where *e* is a binary error vector of length *N* and weight t' and *m* is a binary vector of length *K*.

Decryption

When Alice receives the encrypted message *c* from Bob, she can decipher it as flows:

- $r = cQ = mSG + eQ,$ such that the weight of eQ is at most jt'.
- Using the successive cancellation list decoding algorithm [29] of the private code, she obtains $c' = mS$.
- $m = c'S^{-1}$

Security analysis

In this section we analyse the security of our proposed approach against the most known attacks, structural attacks and generic attacks. Our cryptosystem is able to resist against these attacks.

Structural attacks

Brute force attack

The goal of brute force attack is to recover the secret key from the public key. Hence, this is infeasible when K and N are chosen to be large enough. Since the required number of operations to recover S is $2^{K^2}\Pi_{i=1}^{K}(1-2^i) > 2^{K^2}$ and the number of candidate invertible sparse matrices Q of row and column weight j over $\mathbb{F}_2$ is $\Pi_{i=0}^{N-j}\binom{N-i}{j}$. If $j=1$ then $\Pi_{i=0}^{N-j}\binom{N-i}{j} = N!$ which represents the number of the possible permutation matrix of length N. However in our proposal $j>1$ and thus

$$\prod_{i=0}^{N-j}\binom{N-i}{j} > N!.$$

For the suggested parameters (4096, 1024), $2^{1024^2}\Pi_{i=1}^{1024}(1-2^i) > 2^{1024^2}$ operations to find S, and $\Pi_{i=0}^{4096-j}\binom{4096-i}{j} > 4096!$ Operations to find Q.

Key recovery attack

Key recovery attack is one of the structural attacks, which aims to distinguish the private code from a random code. The most efficient structural attack against the McEliece's cryptosystem is the support splitting algorithm (SSA) proposed by N. Sendrier in 2000 [23] which solve the permutation equivalence problem in polynomial time. Despite the strength of this algorithm, it cannot be effective in the case of polar code since the latter has a large permutation group and large hull. However, in

our proposal this is infeasible since the private code and the public code are not-permutation equivalent and the weight enumerator in both codes not to be the same. However, in [4] M. Bardet et al. succeeded to solve the permutation equivalence problem and break the cryptosystem proposed in [24] where the public code is a permuted version of the private code. The idea of this attack is based on the known structure of low weight codewords of polar code, which can be found easily since polar code is a decreasing monomial code, and by finding the minimum-weight codewords in the public code using the *ISD* attack like Stern [27] or Lion [15] algorithms. However, in our variant, multiplying the generator matrix G of the private code by Q^{-1} which is a dense matrix leads to hide the minimum-weight codewords in the public code that makes searching for them difficult when N is large enough. On the other hand, this is infeasible since Q^{-1} is not a permutation matrix.

Generic attacks

Message recovery attack

In this section, we represent the security analysis against message recovery attack. In this case, the attacker tries to reconstruct the plain-text knowing just its own cipher-text. The most efficient message recovery attack was proposed in the first time by McEliece [17] against its cryptosystem. The goal of this attack is to pick a set of K bits from the cipher-text hope that is error-free. The common name of this attack is Information Set Decoding attack. Several implementations are proposed to improve the cost of this algorithm. Stern in [27] proposed a method that aims to find the small weight codeword in an arbitrary code. In the case of McEliece cryptosystem, the algorithm is applied to the extended code described by the generator matrix $\begin{bmatrix} G \\ c \end{bmatrix}$ where c is the cipher-text and this is equivalent to find the error vector e. We consider the Stern attack [27] to analyse the strength of our proposal. The total work factor of Stern's algorithm required to find a binary vector of weight t' can be expressed by the formula (4), such that p and l are two integers, $0 \le p \le t'$, $0 \le l \le n-k$.

$$WF_{isd} = \left(\frac{\binom{n}{t'}}{\binom{k/2}{p}^2 \binom{n-k-l}{t'-2p}} \right) \left(\frac{(n-k)^2(n+k)}{2} + 2\binom{k/2}{p} pl + \frac{2p(n-k)\binom{k/2}{p}^2}{2^l} \right) \tag{4}$$

Table 1

The error correcting capability of polar code over Binary Symmetric Channel using SCL algorithm with $L = 16$

Rate	(n, k)	t
1/4	(4096, 1024)	590
1/3	(4096, 1365)	450
3/7	(4096, 1750)	284
1/2	(4096, 2048)	194
2/3	(4096, 2730)	80

Table 2

Proposed parameters for 80 bit, 100 bit and 128 bit security levels

Security level	(n, k)	(t', j)	key size (Kbytes)
80 bit	(4096,2730)	(40, 2)	455.22
100 bit	(4096,2048)	(97, 2)	512
128 bit	(4096,1750)	(142, 2)	501.15

Tables 1 and 2 show the error correcting capability of polar code over BSC channel for rates 1/4, 1/3, 3/7, 1/2 , 2/3 using SCL algorithm, and the proposed parameters for 80 bit, 100 bit and 128 bit security levels respectively.

Conclusion

In this paper, we proposed a new McEliece public-key cryptosystem which based on non-permutation equivalent polar code. Since polar code is a family of a decreasing monomial codes any permutation equivalent of the standard polar code can be found due to a powerful structural attack, and any McEliece cryptosystem based on this problem is insecure. However, our proposed cryptosystem is able to resist against this attack and resist against the most known generic attacks.

References

[1] Arikan, E.: Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels. *IEEE Transactions on Information Theory.* 55, 7, 3051–3073 (2009). https://doi.org/10.1109/TIT.2009.2021379.

[2] Baldi, M. et al.: Quasi-Cyclic Low-Density Parity-Check Codes in the McEliece Cryptosystem. In: *2007 IEEE International Conference on Communications.* pp. 951–956 (2007). https://doi.org/10.1109/ICC.2007.161.

[3] Baldi, M., Chiaraluce, F.: Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC Codes. In: *2007 IEEE International Symposium on Information Theory*. pp. 2591–2595 (2007). https://doi.org/10.1109/ISIT.2007.4557609.

[4] Bardet, M. et al.: Cryptanalysis of the McEliece Public Key Cryptosystem Based on Polar Codes. In: Takagi, T. (ed.) Post-Quantum Cryptography. pp. 118–143 Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-29360-8_9.

[5] Berlekamp, E. et al.: On the inherent intractability of certain coding problems (Corresp.). *IEEE Transactions on Information Theory.* 24, 3, 384–386 (1978). https://doi.org/10.1109/TIT.1978.1055873.

[6] Canteaut, A., Chabaud, F.: A new algorithm for finding minimum-weight words in a linear code: application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511. *IEEE Transactions on Information Theory.* 44, 1, 367–378 (1998). https://doi.org/10.1109/18.651067.

[7] Diffie, W., Hellman, M.: New directions in cryptography. *IEEE Transactions on Information Theory.* 22, 6, 644–654 (1976). https://doi.org/10.1109/TIT.1976.1055638.

[8] Dumer, I.: On minimum distance decoding of linear codes. In: Proc. 5th Joint Soviet-Swedish Int. Workshop Inform. Theory. pp. 50–52 (1991).

[9] Faure, C., Minder, L.: Cryptanalysis of the McEliece cryptosystem over hyperelliptic codes. In Proceedings of the 11th international workshop on Algebraic and Combinatorial Coding Theory, ACCT. pp. 99–107 (2008).

[10] Hooshmand, R. et al.: PKC-PC: A variant of the McEliece public-key cryptosystem based on polar codes. *IET Communications.* 14, 12, 1883–1893 (2020). https://doi.org/10.1049/iet-com.2019.0689.

[11] Hooshmand, R. et al.: Reducing the key length of mceliece cryptosystem using polar codes. In: 2014 11th International ISC Conference on Information Security and Cryptology. pp. 104–108 (2014). https://doi.org/10.1109/ISCISC.2014.6994031.

[12] Janwa, H., Moreno, O.: McEliece Public Key Cryptosystems Using Algebraic-Geometric Codes. Designs, Codes and Cryptography. 8, 3, 293–307 (1996). https://doi.org/10.1023/A:1027351723034.

[13] Kakelli Anil Kumar, Addepalli V. N. Krishna & K. Shahu Chatrapati (2017) New secure routing protocol with elliptic curve cryptography for military heterogeneous wireless sensor networks, *Journal of Information and Optimization Sciences*, 38:2, 341-365, DOI: 10.1080/02522667.2016.1220092

[14] Lee, P.J., Brickell, E.F.: An Observation on the Security of McEliece's Public-Key Cryptosystem. In: Barstow, D. et al. (eds.) Advances in Cryptology — EUROCRYPT '88. pp. 275–280 Springer, Berlin, Heidelberg (1988). https://doi.org/10.1007/3-540-45961-8_25.

[15] Leon, J.S.: A probabilistic algorithm for computing minimum weights of large error-correcting codes. *IEEE Transactions on Information Theory.* 34, 5, 1354–1359 (1988). https://doi.org/10.1109/18.21270.

[16] Liu, J. et al.: polarRLCE: A New Code-Based Cryptosystem Using Polar Codes, https://www.hindawi.com/journals/scn/2019/3086975/, last accessed 2020/10/07. https://doi.org/10.1155/2019/3086975.

[17] McEliece, R.J.: A public-key cryptosystem based on algebraic coding theory. DSN Progress Report 42-44. 114–116 (1978).

[18] Minder, L., Shokrollahi, A.: Cryptanalysis of the Sidelnikov Cryptosystem. In: Naor, M. (ed.) Advances in Cryptology - EUROCRYPT 2007. pp. 347–360 Springer, Berlin, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72540-4_20.

[19] Mori, R., Tanaka, T.: Performance of Polar Codes with the Construction using Density Evolution. *IEEE Communications Letters.* 13, 7, 519–521 (2009). https://doi.org/10.1109/LCOMM.2009.090428.

[20] Niederreiter, H.: Knapsack-type cryptosystems and algebraic coding theory. Prob. Contr. Inform. Theory. 15, 2, 159–166 (1986).

[21] Niu, K., Chen, K.: CRC-Aided Decoding of Polar Codes. I*EEE Communications Letters.* 16, 10, 1668–1671 (2012). https://doi.org/10.1109/LCOMM.2012.090312.121501.

[22] Rivest, R.L. et al.: A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM.* 21, 2, 120–126 (1978). https://doi.org/10.1145/359340.359342.

[23] Sendrier, N.: Finding the permutation between equivalent linear codes: the support splitting algorithm. *IEEE Transactions on Information Theory.* 46, 4, 1193–1203 (2000). https://doi.org/10.1109/18.850662.

[24] Shrestha, S.R., Kim, Y.-S.: New McEliece cryptosystem based on polar codes as a candidate for post-quantum cryptography. In: 2014 14th *International Symposium on Communications and Information Technologies* (ISCIT). pp. 368–372 (2014). https://doi.org/10.1109/ISCIT.2014.7011934.

[25] Sidelnikov, V.M.: A public-key cryptosystem based on binary Reed-Muller codes. *Discrete Mathematics and Applications.* 4, 3, 191–208 (1994). https://doi.org/10.1515/dma.1994.4.3.191.

[26] Sidelnikov, V.M., Shestakov, S.O.: On insecurity of cryptosystems based on generalized Reed-Solomon codes. *Discrete Mathematics and Applications.* 2, 4, 439–444 (1992). https://doi.org/10.1515/dma.1992.2.4.439.

[27] Stern, J.: A method for finding codewords of small weight. In: Cohen, G. and Wolfmann, *J. (eds.) Coding Theory and Applications*. pp. 106–113 Springer, Berlin, Heidelberg (1989). https://doi.org/10.1007/BFb0019850.

[28] Sumit Singh Dhanda, Brahmjit Singh & Poonam Jindal (2020) Demystifying elliptic curve cryptography: Curve selection, implementation and countermeasures to attacks, *Journal of Interdisciplinary Mathematics,* 23:2, 463-470, DOI: 10.1080/09720502.2020.1731959.

[29] Tal, I., Vardy, A.: List Decoding of Polar Codes. *IEEE Transactions on Information Theory.* 61, 5, 2213–2226 (2015). https://doi.org/10.1109/TIT.2015.2410251.

Received October, 2020

Revised February, 2021