

The capacity of ciphers fulfilling the accessibility of cryptograms

TOMASZ HEBISZ, EUGENIUSZ KURIATA

University of Zielona Góra, Institute of Control and Computation Engineering

ul. Podgórna 50, 65-246 Zielona Góra, Poland

e-mail: {T.Hebisz, E.Kuriata}@issi.uz.zgora.pl

Abstract: The attempt of using the techniques of error correction coding for building the cryptographic system, which can detect the manipulations on cryptograms, is shown in the paper. Presented approach to generating cipher, generating redundant ciphertexts, which are resistant to manipulations, allows to fulfilling the accessibility as well as confidentiality and authenticity. The capacity of obtained ciphertexts, by mean of statistical tests' results, is also presented.

Key words: cryptography, error correction, security, accessibility of information

INTRODUCTION

The security is defined as fulfillment of the confidentiality, authenticity and accessibility. Most cryptosystems fulfill the confidentiality and authenticity, but ignore or even completely neglect the property of accessibility. There is possible to construct the cryptosystem that gives the accessibility by means of manipulations correction using error correction coding [1, 2].

Error correction codes and cryptography are connected very close and have many common applications. There are a lot of cryptosystems using error correction codes, for example McEliece cryptosystem or Niederreiter cryptosystem [8, 9, 10, 15].

The manipulations on cryptograms are defined as a change of cryptogram contents made by opponent in communication channel. Generally, there is no difference between the transmission errors and manipulations. This justifies the using error correction coding for construction of cryptographic system.

There is possible to apply the non-systematic codes in cryptosystem, which fulfills the accessibility of information. Various modes of operation of such system, and the statistical analysis of cryptograms generated by it, are presented.

1. CRYPTOGRAPHIC ALGORITHM

Proposed cryptosystem de facto generates blocks of ciphertexts, which are vectors of cyclic, non-systematic Reed-Solomon code, and correctional properties of cipher are connected with the parameters of used code.

1.1 Encryption and decryption procedures

Presented cryptographic algorithm is based on using coding and decoding procedures of cyclic Reed-Solomon code. These procedures can be applied for constructing the symmetric-key block cryptosystem, consisting of the pair of transformations $\{E_K, D_K: K \in \mathbf{K}\}$, where $E_K: \mathbf{M} \rightarrow \mathbf{C}$, $D_K: \mathbf{C} \rightarrow \mathbf{M}$ such that

$$(\forall K \in \mathbf{K}, \forall M \in \mathbf{M}: C = E(M + K))$$

and

$$(\forall K \in \mathbf{K}, \forall C \in \mathbf{C}: M = D(C) - K).$$

E denotes encoding procedure of non-systematic cyclic Reed-Solomon code over $GF(q)$, consisting in computing n -symbol code vector C , representing k -symbol information vector $M + K$.

Similarly, D denotes decoding procedure of this code, therefore $D(C) = D(E(M + K)) = M + K$. The symbols $+$ and $-$ denote addition and subtraction in $GF(q)$, respectively. The Fig. 1. shows the block scheme of such system.

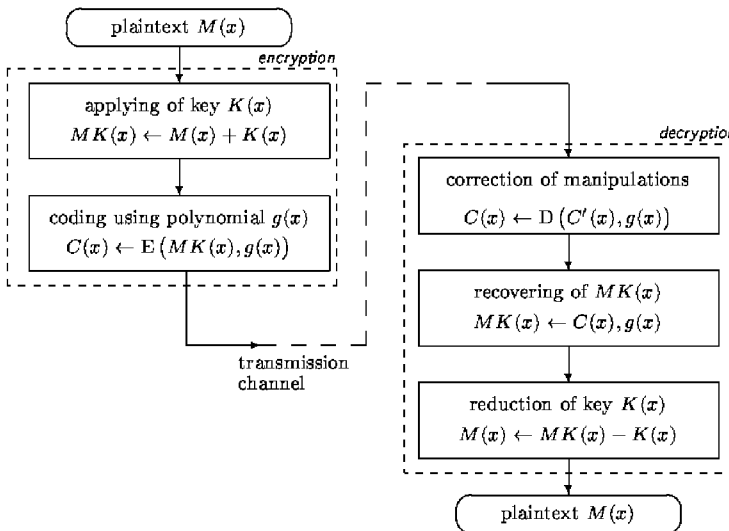


Fig. 1. Block scheme of cryptosystem with using Reed-Solomon code

1.2 Computing in isomorphic finite fields

As the matter of fact, during the encryption procedure the information is coded using nonsystematic Reed-Solomon code. Therefore, the use of different representations of elements of $GF(p^m)$ instead of standard finite field arithmetic devices or routines in the implementations of cryptographic algorithms makes the cryptanalytic work much more difficult. Several essential methods of computing in isomorphic finite fields may be shown [6, 7].

- **“Natural” isomorphism**

This possibility of computing in isomorphic Galois field consists in using different irreducible polynomials while constructing a field. This approach gives

$$N_q(m) = \frac{1}{m} \sum_{e|m} p^{m/e} \mu(e),$$

different representations of elements of $GF(p^m)$ [11]. The symbol $\mu(x)$ denotes the discrete Möbius function described by the following equation

$$\mu(x) = \begin{cases} 1 & \text{if } x = 1, \\ (-1)^k & \text{if } x \text{ is product of } k \text{ different primes,} \\ 0 & \text{if } x \text{ is divided by the square of prime.} \end{cases}$$

- **The application of affined transformations**

Since $GF(p^m)$ is an m -dimensional vector space over $GF(p)$, then any element of $GF(p^m)$ represented using canonical basis can be transformed into the other vector by means of an affined transformation. In this case

$$N_T = p^m \cdot q^{(m-1)m/2} \prod_{i=1}^m q^i - 1$$

isomorphic fields $GF(p^m)$ can be constructed.

- **The application of “general” isomorphism**

This approach uses an arbitrary permutation P mapping elements of $GF(p^m)$ onto the elements in the isomorphic $GF(p^m)$. This method is rather practical for software implementation of operation in finite fields of smaller order gives $(p^m)!$ isomorphic representations of elements.

The presented methods can be easily applied for fast software encryption and in constructing the cryptographic hardware, offering in result more secure level of privacy at some expense of speed of operations.

1.3 Modes of operation of cipher

For proposed cipher two modes of operation are possible: electronic codebook (ECB) or cipher-block chaining (CBC). Algorithm of ECB mode of operation is shown on Fig. 2. It consists of two procedures:

- **Encryption:**

Input: k -symbol key K , k -symbol plaintext blocks M_1, M_2, \dots, M_s .

for $1 \leq i \leq s$ $C_i \leftarrow E(K + M_i)$.

Output: n -symbol ciphertext blocks C_1, C_2, \dots, C_s .

- **Decryption:**

Input: k -symbol key K , n -symbol received ciphertext blocks C'_1, C'_2, \dots, C'_s , certain blocks may have some symbols changed by an intruder or by channel noise.

for $1 \leq i \leq s$

$M_i \leftarrow D(C'_i) - K$.

Output: k -symbol plaintext blocks M_1, M_2, \dots, M_s .

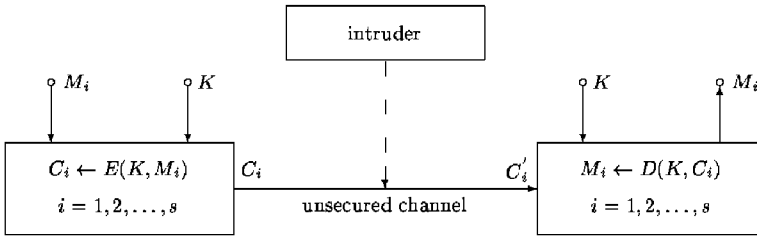


Fig. 2. ECB mode of operation of cipher

This is a simplest approach, where a message is partitioned into k -symbol blocks and each block is encrypted individually using the same key. Blocks are encrypted independently of other blocks. Assuming that proposed cipher can correct t symbols in one block, a change of t or less symbols in a single ciphertext block does not affect decipherment of that block at all. Only when the number of manipulations on the symbols of ciphertext is greater than t and when these manipulations are detectable, deciphering procedure does not work properly, and this fact may be signaled.

Rearranging ciphertext blocks in the channel results in the correspondingly re-ordered plaintext blocks at the receiver. Moreover, identical plaintext blocks are encrypted into identical ciphertext blocks. Ciphertext blocks are independent; therefore, malicious substitution of ciphertext blocks by the intruder does not affect the decryption of adjacent blocks. For these reasons the ECB mode is recommended for encrypting data no longer than one block.

The CBC mode of operation of presented cipher presented on Fig. 3. is more complicated and offers more new possibilities than ciphers with the same length of ciphertext and plaintext, since it is quite different. In comparison with ECB mode,

the CBC mode applies additional elements, namely, memory elements symbolized by squares, in which one n -component vector over $GF(q)$, representing one block of ciphertext, can be stored, and a block, performing the operation of function f_K . The cryptographic procedures in this mode of operation result directly from Fig. 3. and are the following:

- **Encryption:**

Input: secret function f_K , n -component initialization vector (I_V), k -symbol plaintext blocks M_1, M_2, \dots, M_s .

$$C_0 \leftarrow I_V,$$

for $1 \leq i \leq s$

$$\begin{aligned} K_i &\leftarrow f_K(C_{i-1}), \\ C_i &\leftarrow E(K_i + M_i). \end{aligned}$$

Output: n -symbol ciphertext blocks C_1, C_2, \dots, C_s .

- **Decryption:**

Input: secret function f_K , n -component initialization vector (I_V), n -symbol received ciphertext blocks C'_1, C'_2, \dots, C'_s , certain blocks may have some symbols changed by an intruder or by channel noise.

$$C_0 \leftarrow I_V,$$

for $1 \leq i \leq s$

$$\begin{aligned} K_i &\leftarrow f_K(C_{i-1}), \\ C_i &\leftarrow D^*(C'_i), \\ M_i &\leftarrow D(C'_i) - K_i. \end{aligned}$$

Output: k -symbol plaintext blocks M_1, M_2, \dots, M_s .

It should be noted that D^* function carries out the correction of possible manipulations on C'_i and must always be calculated, as an intermediate stage of decrypting procedure D.

Since CBC mode is more sophisticated than ECB mode, its properties are also more rich than those of the latter. When the same plaintext of s blocks is enciphered under the same key and initialization vector, identical ciphertext blocks will be obtained. If one wants to obtain different ciphertexts for the same message, encrypted two times, he must change I_V or first plaintext block.

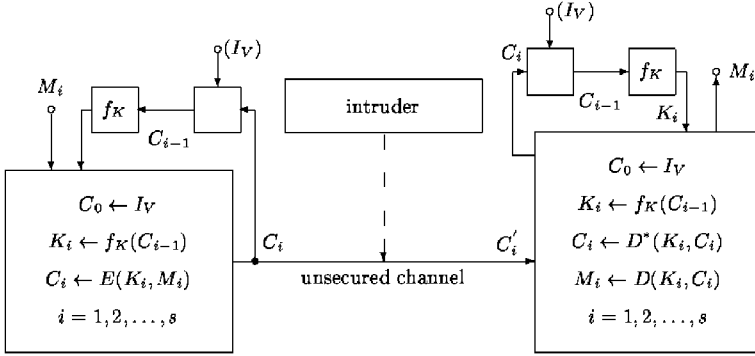


Fig. 3. CBC mode of operation of cipher

The application of chaining mechanism makes the ciphertext C_i dependent on M_i and all preceding plaintext blocks. Rearrangement of the order of ciphertext blocks in the channel by an intruder causes then mistakes in decryption. Correct decryption of block C_i requires a proper C_{i-1} block.

It follows from the properties of shown cipher that after changing in ciphertext C_i more than t symbols, the decipherment of block C_i and C_{i+1} will be erroneous.

The CBC mode is self-synchronizing in the sense that if in ciphertext block C_v the manipulations cannot be corrected but C_{v+1} has less than $(t + 1)$ manipulations, then ciphertext block C_{v+2} is correctly decrypted to M_{v+2} .

2. ANALYSIS OF CIPHERTEXTS

The cryptographic hardness of proposed method depends to a large degree of the key space size. Generally, in presented method the message is coded using redundant code. Thus the appreciating of cryptographic strongness of ciphertexts is verified by analysis of three problems:

- computation complexity of analytic cryptanalysis,
- defining the key space,
- analysis of statistical properties of ciphertexts.

In presented cryptosystem, first criterion is based on complexity of decoding cyclic codes without the knowledge of generator polynomial. This approach is based on syndrome decoding problem and it has successfully resisted more than 20 years of cryptanalysis effort, because it is still NP-problem.

The key space is depending on the way of computation in isomorphic finite field. For example, for “general” isomorphism, the size of key space consists of:

- number of polynomial $K(x)$ added to the message, which is equal to

$$(p^m)^k$$

where (p^m) — size of used GF , k — number of information symbols in codeword,

- number of permutations used for generate isomorphic Galois fields applied for coding/decoding procedure, which is equal to

$$(p^m)!$$

Thus, for established parameters of code, such as p , m , n and k ,

$$|\mathbf{K}| = (p^m)^k ((p^m)!).$$

For example, let us use the $GF(2^8)$, because of very huge popularity of the *ASCII* alphabet, and let $k = 100$, $n = 257$ ($t = \lfloor 257-100 \rfloor / 2 = 78$), we obtain:

$$|\mathbf{K}| = (2^8)^{100} (2^8)! = 256^{100} (256!)$$

For analysis of statistical properties of ciphertexts one can use the statistical tests and histograms of character occurrence frequency in ciphertext.

During the computer experiment, many attempts of ciphering were performed. A task of computer experiments was to prove the correction advantages of the system. A text file and an audio-wave file were used as plaintext files, because of its characteristic frequency of character occurrence. The size of used files was about 8MB. The cryptograms files were bigger then the plaintexts files and they took about 11 MB space of hard disk.

Fig. 4. presents an example of the histograms of character occurrence frequency in plaintexts and ciphertexts obtained during enciphering text file and wave-audio file. As one can see, the characters of ciphertexts have almost proportional distribution. Much better results give the CBC mode of cipher. A cryptogram resembles the random set of characters.

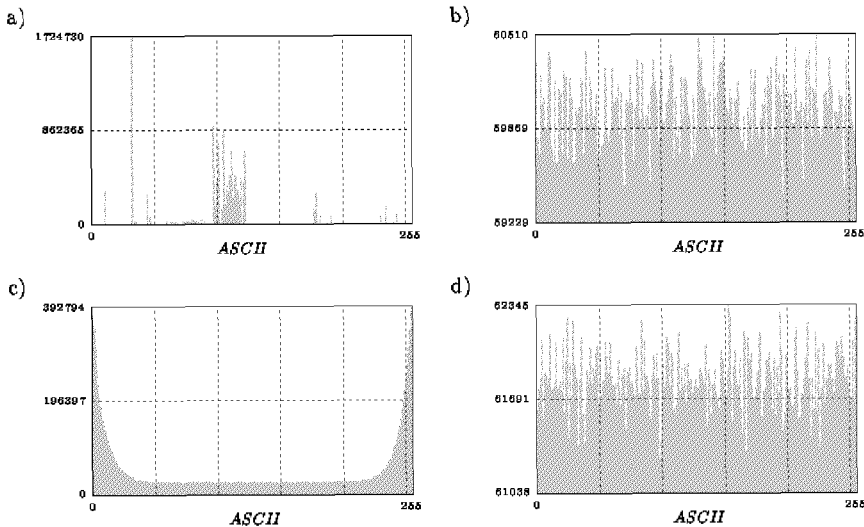


Fig. 4. Histograms of character occurrence frequency a) in text file, b) in ciphertext of text file, c) in wave-audio file and d) in cipher-text of wave-audio file

Test	Max	a	b	c	d	e
BIRTHDAY SPACINGS	1	1	1	1	1	1
THE OVERLAPPING 5-PERMUT.	2	0	2	1	2	1
BINARY RANK for 31×31 mat.	1	1	1	1	1	1
BINARY RANK for 32×32 mat.	1	1	1	1	1	1
BINARY RANK for 6×8 mat.	1	1	1	1	1	1
THE BITSTREAM TEST	20	0	20	16	19	20
OPSO	23	0	22	22	22	22
OQSO	28	0	27	28	28	28
DNA	31	0	30	30	30	30
COUNT-THE-1's on a bytes stream	2	0	2	1	2	2
COUNT-THE-1's for spec. bytes	25	22	25	24	25	22
THE PARKING LOT	1	1	1	1	0	1
THE 3DSPHERES	1	0	1	1	1	1
SQUEEZE	1	1	1	1	1	1
THE OVERLAPPING SUMS	1	0	1	1	1	0
THE RUNS	4	4	4	4	3	4
THE CRAPS	1	1	1	1	1	1

Tab. 1. Results of DIEHARD tests of cryptograms

- a) cryptograms for text file generated in ECB mode, b) cryptograms for text file generated in CBC mode, c) cryptograms for wave-audio file generated in ECB mode, d) cryptograms for wave-audio file generated in CBC mode, e) cryptograms for text file encrypted using T-DES algorithm in ECB mode.

There are few packets of statistical tests available for cryptographic applications. The Diehard tests packet appears to be most popular [12]. Tab. 1. shows the results of Diehard's tests of cryptograms obtained for enciphered textfile and wave-audio file.

Unfortunately there is not the measure which allows to reliably compare of statistical properties of various cryptographic algorithms. The only possibility is to compare the results of individual tests results. To compare obtained results with other cryptosystems, the text file used in the above example was encrypted using the T-DES method which is often applied in practice. Fig. 5. shows the histograms of character occurrence frequency and the Tab. 1(e) shows the results of Diehard's tests for the obtained ciphertext. As one can see, the obtained results for the method T-DES is similar to the results of case using Reed-Solomon code.

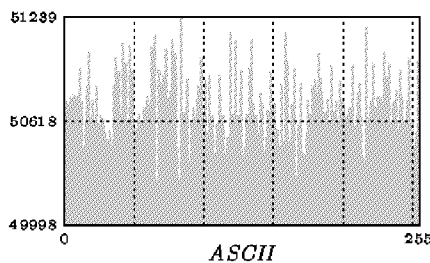


Fig. 5. Histogram of character occurrence frequency for ciphertext obtained using T-DES algorithm

3. CONCLUSIONS

Presented cryptosystem allows verifying the correctness of obtained codewords, what gives an opportunity to fulfill the confidentiality, authenticity and accessibility, using the properties of cyclic Reed-Solomon code.

Unfortunately there is no measure, which allows to reliably comparing of statistical properties of various cryptographic algorithms. But in consideration of Diehard tests' results and histograms of character occurrence frequency of obtained ciphertexts, one can say, that the statistical properties of ciphertexts are sufficient and hard enough to cryptanalysis. Long enough key space is guaranteed by using computation in isomorphic GF . Especially worth mentioning is CBC mode of shown cipher. A cryptograms produced in this mode have very good statistical properties however the computational complexity is bigger than in ECB mode, but the difference seems to be insignificant.

The presented method are equally suitable for the software and hardware implementations as well, but the hardware implementation seems to has much more practical applications for the sake of speed of ciphering and deciphering procedures.

REFERENCES

- [1] T. Hebisz, Cz. Kościelny. *A method of constructing symmetric-key block cryptosystem resistant to manipulations on ciphertext*. Bulletin of the Polish Academy of Sciences, Technical Sciences, Vol. 50, No. 4, 2002.
- [2] T. Hebisz, E. Kuriata, M. Jackiewicz. *Fulfilment of computer security and safety by using symmetric-key block cryptosystem resistant to manipulations on ciphertext*. International Conference on Computer Information Systems and Industrial Management Applications CISIM '03, 2003.
- [3] A. Kiayias, M. Yung *Polynomial Reconstruction Based Cryptography*. SAC 2001. ICALP 2002. LNCS 2259. pp. 129-133. Springer-Verlag, 2002.
- [4] A. Kiayias, M. Yung *Cryptographic Hardness Based on the Decoding of Reed-Solomon Codes*. Springer-Verlag. ICALP 2002. LNCS 2380. pp. 232-243. 2002.
- [5] L. Knudsen, B. Preneel *Construction of Secure and Fast Hash Function Using Nonbinary Error-Correcting Codes*. IEEE Trans. on Information Theory. Vol. 48. No. 9. pp. 2524-2537. 2002.
- [6] Cz. Kościelny. *Computing in the composite $GF(q^m)$ of characteristic 2 formed by means of an irreducible binomial*, International Journal of Applied Mathematics and Computer Science, Vol. 8, No. 3, pp. 671-680, 1998.
- [7] Cz. Kościelny, T. Hebisz. *More secure computing in finite fields for cryptographic applications*. Mathematical Theory of Networks and Systems MTNS 2000, The fourteenth International Conference, Perpignan, 2000, CD-ROM.
- [8] E. Krouk. *A new Public Key Cryptosystem*. Proc. of Sixth Joint Swedish-Ruppian Intern. Workshop on Information Theory, 1993.
- [9] E. Kuriata. *Error correction codes in cryptography*. VI Intern. conference "Wojskowa Konferencja Telekomunikacji i Informatyki", 1997 (in polish).
- [10] Y. X. Li, R. H. Deng, X. M. Wang. *On the equivalence of McEliece's and Niederreiter's public-key cryptosystems*. IEEE Trans. on Information Theory. Vol. 40. pp. 271-273. 1994

- [11] R. Lidl, H. Niederreiter. *Introduction to finite fields and their applications*. Cambridge University Press, 1986.
- [12] G. Marsaglia. *Statistical tests Diehard*. <http://stat.fsu.edu/~geo/diehard.html>.
- [13] R. J. McEliece. *A Public Key Cryptosystem Based on Algebraic Coding Theory*. JPLDSN Progrepp Rept., pp. 42-44, 1978.
- [14] A. J. Menezes, ed. *Application of Finite Fields*. Kluwer Academic Publishers, 1993.
- [15] H. Niederreiter. *Knapsak-type cryptosystems and algebraic coding theory*, Probl. Control and Inform. Theory, Vol. 15, 1986.

Enhanced Methods in Computer Security, Biometric and
Artificial Intelligence Systems

Pejas, J.; Piegat, A. (Eds.)

2005, XII, 396 p., Hardcover

ISBN: 978-1-4020-7776-0