UDC 681.3.06

*S. Yevseiev, DSc (Simon Kuznets Kharkiv National University of Economics, Ukraine),*
*R. Korolev, PhD (Kharkov National Air Force University, Ukraine)*

# Mathematical model of hybrid crypto code constructions on damaged codes

*The subject of the article is a formal mathematical description of the developed hybrid crypto-code constructions on defective codes. The goal is to secure security requirements for 2FA based on OTP passwords in banking systems. The object is mathematical model for the implementation of hybrid crypto-code constructions on defective codes.*

Automated banking systems (ABS) are increasingly using the global Internet (GIS) and its main portal – mobile communications to provide the full range of financial services, electronic document management, and administrative functions. One of the main components of security when using different technologies and gadgets is electronic authentication (EA) – a procedure that confirms the authenticity of the source of the message. The downside of using OTP passwords is the ability to “intercept” an attacker's text (SMS) with one part of the token. Attackers can compromise your two-factor authentication based on the text in several ways: based on social engineering techniques (message forwarding through the provider), intercept messages using the International Mobile Subscriber Identity (IMSI), use of shortcomings in protocols that allow Operators to exchange data between networks [15, 16]. For this reason, the National Institute of Standards and Technology (NIST) in [6] is ready to prohibit the use of two-factor authentication codes based on OTP passwords for services that connect to public IT systems. Thus, there is a contradiction between the use of OTP passwords in the protocols of two-factor authentication and security in the transmission of its individual factors.

**Basic principles of constructing cryptosystems on defective codes.** In [24, 25], theoretical and practical bases of construction of defective codes are considered. By *flawed text is understood the text obtained by further deformation of non-redundant codes of letters.*Thus, a necessary and sufficient condition for the loss of text with loss of meaning is the shortening of the lengths of the code symbols of the text beyond their redundancy. As a consequence, the defective text has a length shorter than the length of the source text, and there is no sense in the source text [24].

The theoretical basis for the construction of defective texts is to remove the ordering of the symbols of the source text and, as a consequence, to reduce the redundancy of the language symbols in the flawed text. In this case, the amount of information expressing this ordering will be equal to the decrease of the entropy of the text in comparison with the maximum possible entropy value corresponding to the lack of ordering in the text in general, i.e. Equiprobably appearance of any letter after any previous letter. The methods of computing information proposed by K. Shannon allow us to determine the ratio of the amount of predictable information (formed according to certain rules) and the amount of that unexpected information that cannot be predicted in advance. Redundancy of the text is calculated by the formula:

$$B(M) = B_A L_0 = \left( \log N - \frac{H(M)}{L_0} \right) \times L_0 ,$$

where *M* – original text; *B* – language redundancy ( $B = R - r$ , *R* – absolute entropy of a language ( $R = \log N$ , *N* – alphabet power, *r* – language entropy per character, $r = H(M)/L$ , *L* – length of the *M* message in the language symbols); *H(M)* – entropy (uncertainty) of message; $L_0$ – the length of the message *M* in the symbols of the language with meaning; $B_A$ – redundancy of the language.

To obtain a defective text (FTC) and damage (DCH), the "ideal" compression method is used after completing the *m* cycles of the damage mechanism $C_m$ [24, 25].

The number of cycles required to reduce the length of the source text is:

$$m \rangle \frac{\log n - B_A}{\log \eta},$$

where *n* – the power of representing the character of the source text; $B_A$ – language redundancy; $\eta$ – the number of times the original text length in MV2 decreases at each step (some constant coefficient).

A quantitative measure of the effectiveness of damage is the degree of destruction of the meaning, equal to the difference in the entropies of the defective text and the source text at different lengths of the defective text:

$$d = H(FTC) - \sum_{i=1}^{s} H(M_i)p_i, \quad \sum_{i=1}^{s} p_i = 1, \quad s = \left[ \frac{L_0 - L_{FTC}}{L_{FTC}} \right],$$

where $M_i$ – part of the source text corresponding to the i-th segment, $p_i$ – its probability, $L_0$ – length $M_i$ equal to the length $L_{FTC}$ – flawed text, *s* – number of segments. For an ergodic source of source code characters:

$$d_{max} = \log L_{FTC} - H(M_i).$$

Under *the information core* of some text is understood the flawed text of CFT, obtained by cyclic transformation of the universal damage mechanism $C_m$.

Universal damage mechanism $C_m$ can be described [24, 25]:

$$CFT / CH_{FT} = E_1 \;\; M, KU^{EC} \;\; , CHD / CH_D = E_2 \;\; M, KU^{EC} \;\; ,$$

$$M = E_{1,2}^{-1}( CFT / CH_{FT}, CHD / CH_D, KU^{EC} ),$$

where

$$CFT / CH_{FT} = CFT / CH_{FT}^{i}, \ldots, CFT / CH_{FT}^{m}, KU^{EC} = \varphi( K_D^{i}, \ldots, K_D^{m}, KU_1^{EC}, \ldots, KU_m^{EC},$$

$$CHD / CH_D = CHD / CH_D^{i}, \ldots, CHD / CH_D^{m} \qquad .$$

Thus, as a result we have two ciphertexts (damage ($CH_D$) and flawed text (*FTC*)), each of which has no meaning either in the source code alphabet or in the alphabet of the ciphertext. Actually, the ciphertext of the original message (*M*) is represented in the form of a set of two defective ciphertexts, each of which cannot individually restore the original text. To restore the original sequence, there is no need to know the intermediate faulty sequences. It is necessary to know only the last flawed sequence (the last flawed text after all the cycles) and all the damages with the rules for their application.

**Mathematical model of McElis on defective codes, practical algorithms for their implementation.**

Consider a formal description of the modified McEliece crypto code system on the defective codes used in the two-factor authentication protocol.

To construct a mathematical model, we use the basic propositions in [25] for a formal mathematical definition of a secret system. In [22], a formal description of the mathematical model of crypto-code constructs on defective codes (GCCCDC) McEliece on modified elliptic codes was considered, in [1] a universal mechanism for causing damage and methods of transmission in systems on defective codes were considered. In GCCCDC McEliece, the modified (truncated) algebraic geometric *(n, k, d)* code $C_{k-h_j}$ With a fast decoding algorithm is masked for random *(n, k, d)* code $C_{k-h_j}$ * by multiplying the generator matrix $G^{EC}$ code $C_{k-h_j}$ on secretive masking matrices $X^u$, $P^u$ and $D^u$, Providing the formation of an authorized user's public key: $G_X^{ECu} = X^u \cdot G^{EC} \cdot P^u \cdot D^u, \quad u \in \{1, 2, ..., s\}$, where $G^{EC}$ – generating $n \times k$ Matrix of algebraic geometric block $(n,\, k,\, d)$ code with elements from $GF(q)$, Constructed on the basis of the use of the coefficients of the polynomial of the curve chosen by the user $a_1 ... a_6$, $\forall a_i \in GF(q)$, uniquely defining a particular set of points of a curve from the space $P^2$.

Formation of closed text $C_j \in C_{k-h_j}$ On the entered open text $M_i \in M$ and the given public key $G_X^{ECu}\ {}_{a_i}$, $u \in \{1, 2, ..., s\}$ Is performed by forming a codeword of the masked code with the addition of a randomly generated vector $e = (e_0, e_1, ..., e_{n-1})$:

$$C_j = \phi_u \;\; M_i, G_X^u \;\; = M_i \cdot \;\; G_X^u \;\; {}^T + e,$$

The Hamming weight (the number of non-zero elements) of the vector $e$ does not exceed the correcting ability of the algebraic block code used:

$0 \le w \;\; e \;\; \le t = \left\lfloor \dfrac{d-1}{2} \right\rfloor$, $\lfloor x \rfloor$ – The integer part of the real number *x*.

For each closed text being generated $C_j \in C_{k-h_j}$ corresponding vector $e = (e_0, e_1, ..., e_{n-1})$ Acts as a one-time session key, i.e. for a specific $E_j$ vector $e$ is formed randomly, equiprobably and independently of other closed texts. The MV2 algorithm receives $C_j^* = C_j - C_{k-h_j}$, $E_{K_{MV2}} : C_j^* \rightarrow \left\| f \;\; x \;\; {}_i \right\| + \left\| C \;\; x \;\; {}_i \right\|$.

In the communication channel $\left\| f \;\; x \;\; {}_i \right\|$ and $\left\| C \;\; x \;\; {}_i \right\|$, In this case, the transmission can be carried out either one at a time or two independent channels.

On the receiving side, an authorized user who knows the rule of damage $F_n^r$, Masking, the number and location of null information symbols can take advantage of the fast algorithm for decoding algebraic geometric code (polynomial complexity) for recovering plain text: $E_{K_{MV2}}^{-1} : \left\| f \;\; x \;\; {}_i \right\| + \left\| C \;\; x \;\; {}_i \right\| \rightarrow C_j^*$, $M_i = \phi_u^{-1} \;\; C_j^*, \{X, P, D\}_u \;\;$.

To recover plaintext, an authorized user adds null information symbols $C_j^* = C_j + C_{k-h_j}$, from recovered private text $C_j$ Removes the action of secret permutational and diagonal matrices $P^u$ and $D^u$ :

$$C = C_j^* \cdot {D^u}^{-1} \cdot {P^u}^{-1} = M_i \cdot {G_X^u}^T + e \cdot {D^u}^{-1} \cdot {P^u}^{-1} = M_i \cdot {X^u \cdot G \cdot P^u \cdot D^u}^T + e \cdot {D^u}^{-1} \cdot {P^u}^{-1} =$$

$$= M_i \cdot {X^u}^T \cdot G^T \cdot {P^u}^T \cdot {D^u}^T \cdot {D^u}^{-1} \cdot {P^u}^{-1} + e \cdot {D^u}^{-1} \cdot {P^u}^{-1} = M_i \cdot {X^u}^T \cdot G^T + e \cdot {D^u}^{-1} \cdot {P^u}^{-1}.$$

Decodes the resulting vector according to Berlekamp-Messi algorithm [15]: $C = M_i \cdot {X^u}^T \cdot {G^{EC}}^T + e \cdot {D^u}^{-1} \cdot {P^u}^{-1}$. Eliminates the second term and the factor $G^{ECT}$ In the first term on the right side of the equation, and then removes the action of the masking matrix $X^u$. To do this, the resulting decoding result $M_i \cdot {X^u}^T$ Should be multiplied by ${X^u}^{-1}$: $M_i \cdot {X^u}^T \cdot {X^u}^{-1} = M_i$. The solution is the essence of plain text $M_i$ .

## Conclusion

A mathematical model of a hybrid crypto-code structure based on the modified McEliece cryptosystem is proposed, which differs from the known scheme by reducing the power of the alphabet without decreasing the cryptographic strength. The use of the initialization vector and modified elliptical truncated codes makes it possible to reduce the energy costs for practical implementation, to provide the required level of cryptographic stability. The conducted researches in the work confirm that their application provides encryption speed at the level of application of block symmetric ciphers, provable cryptostability on the basis of the theoretic-complexity problem of decoding a random code (provided $10^{30}$ – $10^{35}$ group operations), and reliability based on the use of a truncated algebraic geometric code (provided $P_{err}$ $10^{-9}$ – $10^{-12}$). To further reduce the power of the alphabet (the Galois field up to *GF* ($2^4$ – $2^6$) In the work it is proposed to use systems on defective codes, which allow simultaneously to form multi-channel cryptosystems

## References

1. Mishchenko V. A., Vilansky Yu. V. Damage texts and multichannel cryptography. Minsk. Encyclopedic. 2007.

2. Shannon K. E. The theory of communication in secret systems. In the book: Shannon KE. Work on the theory of information and cybernetics. - Moscow: Il. 1963.

3. Evseev S.P. Development of the modified asymmetric crypto-code system of McEliece on truncated elliptic codes. Evseev, Kh. N. Rzaev, O. G. Korol / East European Journal of Advanced Technologies. Kharkiv. 2016. Issue. 4/9 (82). P. 4–12.

4. Scott Rose. Domain name systems-based electronic mail security// [Electronic resource] : – https://nccoe.nist.gov/sites/default/files/library/sp1800/dns-secure-email-sp1800-6-draft.pdf.

5. Security requirements for cryptographic modules // [Electronic resource] : – https://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf.