UDC 004.056.55

KEY SCHEDULE OF BLOCK SYMMETRIC CIPHERS

Alexandr Kuznetsov[1], Yuriy Gorbenko[1], Ievgeniia Kolovanova[1]

V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
kuznetsov@karazin.ua, YuGorbenko@iit.kharkov.ua, e.kolovanova@gmail.com

Reviewer: Victor Dolgov, Dr., Full Professor, V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
dolgovvi@mail.ru

Received on June 2016.

Abstract. We investigate combinatorial properties of the block symmetric ciphers key schedule in the assumption that the cyclic (round) keys are generated randomly, with equal probability and independently of each other. The model of random homogeneous substitution is used for an abstract description of this formation. Analytical expressions allow us to estimate the power of implemented encryption-decryption maps set, obtain estimates of the probability properties of round keys sequences and ratios of the average number of different key sequences to power of different master keys set. The simulation results confirm the accuracy and validity of these analytical expressions.

Key words: key schedule, cyclic keys, combinatorial properties, block symmetric ciphers.

1 Problem statement and analysis of the literature

Ciphering is widely used in modern information and telecommunication systems for information protection and security. Ciphering is a reversible cryptographic transformation of open data to hide its semantic content from unauthorized user (attacker). Bijective processes of encryption and decryption of plaintext blocks and ciphertext blocks are parameterized by key data, which is the same for symmetric cryptographic transformation [1].

Most of block symmetric ciphers (BSC) are iterative [1], so the encryption is realized by cyclically repeating reversible round function (Fig. 1). The round (cyclic) keys $K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}$ are used for parameterization of round transformations at each iteration of BSC. These keys are formed by extending (key scheduling) the master key $K^{(x)}$ [1].

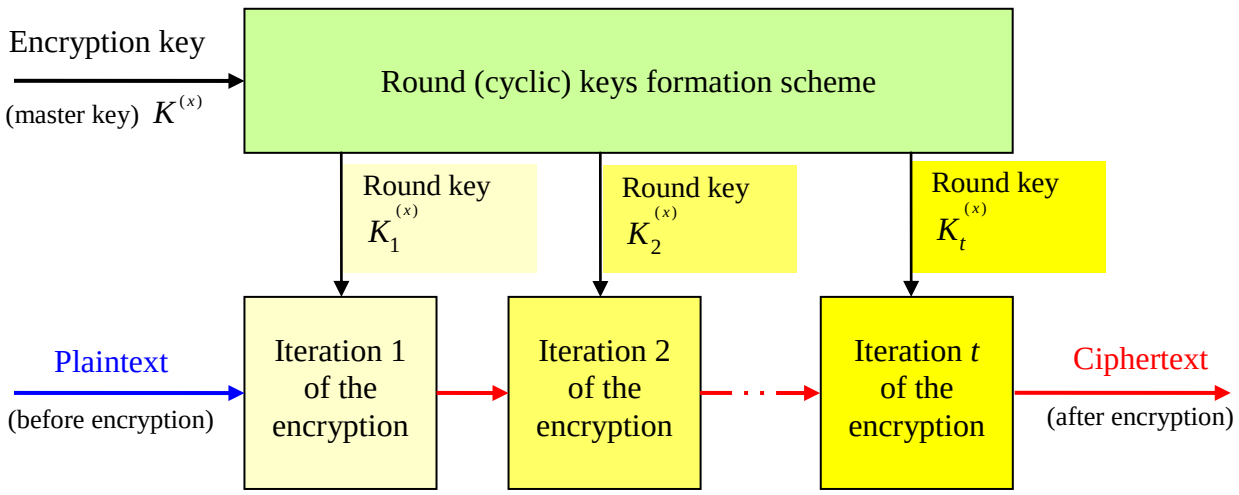

Fig. 1 - Block diagram of an iterative block cipher

The structure of iterative BSC key schedule and simplicity of round keys formation and/or interdependence are used in the known attacks on the key schedule construction, especially slide attack [2-4], related-key attack [5-7], etc [1].

In the simplest case the key schedule construction can consist of master key repetition for each round. A similar approach was used in the formation of cyclic keys in the Soviet symmetric algorithm of cryptographic transformation GOST (State Standard) 28147-89, which is now also the encryption standard of Ukraine DSTU (State Standards of Ukraine) GOST 28147: 2009 [1]. However, in the case where to the input of each round function (see. Fig. 1) a certain key is fed, and this key is the same for all rounds, the cipher becomes vulnerable to slide attack [2,3]. The option when deployment function involves cyclic repetition of a certain set of round keys (round self-similarity ciphers) can also be easily reduced to this case [4].

To confront the key schedule cryptanalytic attacks modern BSC use the complicated round keys schedule construction implemented using conversion cipher transformations. One of these BSC is the US national standard FIPS -197 (AES) [8,9], adopted in 2001. It is an international algorithm, which is the most prevalent in today's security protocols. The key schedule of BSC AES is a linear array of 4-byte words. The first elements of the array contain master encryption key, the rest are determined recursively by modulo summation of two previous items. For certain positions of the array additional cipher transformation is also applied, in particular, the nonlinear permutation data block, and cyclic shift and etc. [8,9]. As a result, a sequence of round key $K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}$ is formed which non-linearly dependent's on the original master key $K^{(x)}$, and this additional nonlinearity can effectively resist slide attacks on key schedule [1].

Related-key attacks were first proposed in [5] and further developed in [6,7]. In particular, the first cryptanalytic attack on the basis of related keys on a full-cipher AES-192 and AES-256 (variants of FIPS-197 with key lengths 192 and 256 bits) was described in [7]. It should be noted that the attacks in [7] are more effective than the full search of master keys, i.e. we can talk with certainty about the actual decrease of standardized cryptographic algorithm resistance.

Thus, the attacks on the key schedule are continuously improved and their possible use represents a real security threat to modern information systems and technologies [1-7]. Efficient BSC must effectively resist to the key schedule attacks and the key schedule construction must not contain any vulnerabilities caused by the simplicity of formation and the mutual dependence of cyclic keys [11]. In fact, we are talking about "ideal" round keys deployment, when each element of the sequence $K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}$ are generated randomly, with equal probability, and independently of the other cyclic keys. Only in this case we can talk with certainty about the futility of the key schedule attacks because each round BSC is parameterized by randomly chosen value and would operate independently from other iterations of the encryption scheme (see. Fig. 1).

As an example of the key schedule schemes development we can use the algorithm "Kalyna", adopted as a national standard of BSC in Ukraine [12]. It has enhanced the cyclic key schedule construction, due to the use of special one-way functions. The cyclic keys of BSC "Kalyna" are formed as a result of several rounds of encryption, parameterized by auxiliary key. The auxiliary key, in its turn, is also formed as a result of multiple rounds encryption parameterized by master key. In other words, the separate elements of the cyclic keys sequence $K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}$ are generated by independent encrypting of various input data blocks on different keys. Assuming that applied encryption implements are a random substitution (permutation) of data blocks [13-14], then the resulting round keys are generated randomly, with equal probability and independently of each other [11,15]. In particular, in [15] the properties of the key schedule of BSC "Kalyna" are investigated to confirm the resilience of the cipher to related-key attacks and attacks on implementation.

It should be noted that, even under random, equiprobable and independent formation of the round key the corresponding sequence can be the same, what is equivalent to reduction the power of encryption-decryption implemented maps set.

The aim of this work is to analyses combinatorial properties of BSC key schedules, provided that cyclic keys are generated randomly, with equal probability and independently of each other. The model of random homogeneous substitution is used for an abstract description of this formation. The practical benefit of this research results consists in providing its interpretation in order to assess the properties of the key schedule in recently adopted national standard BSC of Ukraine.

2 Random substitution as a model for the cyclic keys formation

Let us consider the definition and basic properties of a random substitution (permutation) [13,14], that will be used further to assess the probability properties of round key sequences.

In combinatorics, a permutation is an ordered set of numbers $1,2,\ldots,n$, that is a bijection on the set $\{1,2,\ldots,n\}$, which puts the *i*-th elements of the set in correspondence to the i number. The number n in this case is called the order (degree) of permutation [13,14].

The substitution s of arbitrary set $Y=\{y_1,y_2,...,y_n\}$ is a rule that each element y_i of set Y puts in correspondence some other element $s(y_i)$ [13,14]:

$$s=\begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ s(y_1) & s(y_2) & \ldots & s(y_n) \end{pmatrix}.$$

In group theory the substitution is a bijection of this set into itself, i.e. substitution s degrees n is considered as a permutation of the elements of the set $Y=\{y_1,y_2,...,y_n\}$ and for all $i=1,2,...,n$ corresponding $s(y_i)\in Y$.

The function $s(y_i)$ value for a specific element $y_i\in Y$ will be called the implementation of substitution s in *i*-th point.

The composition of substitutions s_u and s_w degree n is defined as the consistent fulfillment of the set Y elements permutation [13, 14]:

$$s_u\circ s_w=\begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ s_u(s_w(y_1)) & s_u(s_w(y_2)) & \ldots & s_u(s_w(y_n)) \end{pmatrix},\ s_u(s_w(y_i))\in Y.$$

Concerning operations of sequential substitutions execution the set of all $n!$ permutations degree n forms a group, called the symmetric group and denoted as $S_n=\{s_1,s_2,...,s_{n!}\}$.

By definition [13,14], random substitution (permutation) s_x is the random vector $\{s_x(y_1),s_x(y_2),...,s_x(y_n)\}$, where all elements take values from the Y set and the probability of a match of any two elements is equal to 0. In other words, a random substitution is randomly chosen permutation from the set S_n

$$s_x=\begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ s_x(y_1) & s_x(y_2) & \ldots & s_x(y_n) \end{pmatrix},\ s_x\in S_n,\ x\in\{1,2,...,n!\},$$

defined by a set (vector) of random values $\{s_x(y_1),s_x(y_2),...,s_x(y_n)\}$, that match probabilities satisfy the following condition:

$$\forall i\neq j\in\{1,2,...,n\}: P(s_x(y_i)=s_x(y_j))=0.$$

Thus, under the implementation of a random substitution s_x we mean the specific implementation of random vector $\{s_x(y_1),s_x(y_2),...,s_x(y_n)\}$ and the corresponding value of the $s_x(y_i)$ function we will call the implementation of a random substitution s_x in the *i* -th point.

Independent random substitution is called such a random permutation $\{s_x(y_1),s_x(y_2),...,s_x(y_n)\}$, for which is true

$$P(s_x) = \frac{P(s_x(y_1))P(s_x(y_2))...P(s_x(y_n))}{\sum_{u=1}^{n!} P(s_u(y_1))P(s_u(y_2))...P(s_u(y_n))}.$$

If $\forall i,u \in \{1,2,...,n\} : P(s_u(y_i)) = n^{-1}$ then

$$P(s_x) = \frac{P(s_x(y_1))P(s_x(y_2))\dots P(s_x(y_n))}{\sum_{u=1}^{n!} P(s_u(y_1))P(s_u(y_2))\dots P(s_u(y_n))} = \frac{n^{-n}}{\sum_{u=1}^{n!} n^{-n}} = \frac{1}{n!} \quad (1)$$

and s_x is called the random, equiprobable and independent (or, in abbreviated form, homogeneous) random substitution.

Thus, the concept of a random homogeneous substitution corresponds to a uniform probabilistic distribution on the set $S_n = \{s_1, s_2, ..., s_{n!}\}$ with the independent implementation of random vectors

$$\{s_x(y_1), s_x(y_2), ..., s_x(y_n)\} \quad (2)$$

$$s_x \in S_n, \ \forall x \in \{1,2,...,n!\} : P(s_x) = (n!)^{-1}.$$

Modern BSC are commonly described by the random homogeneous substitution model [1,11], i.e. it is a standard assumption that probabilistic properties of a processed data blocks bijection implemented by encryption function, satisfies the characteristics of a random substitution.

Indeed, if random, equiprobable and independent selection of the master key $K^{(x)}$ is associated with the choice of substitution $s_x \in S_n$, then the resulting ciphering transformation will match a random, equiprobable and independent comparison of ciphertext blocks to plaintext blocks on all possible options of bijective mapping, parameterized by key. For instance, for l-bit cipher with a k bit master key the model of random substitution will consist of subset

$$S'_n = \{s'_1, s'_2, ..., s'_{2^k}\} \subset S_n = \{s_1, s_2, ..., s_{n!}\}$$

with random, equiprobable and independent 2^k substitutions degree $n = 2^l$ (acting on the set $Y = \{y_1, y_2, ..., y_{2^l}\}$ of binary data blocks). At that the choice of substitution $s'_x \in S'_n \subset S_n$ (implementation of random vector $\{s_x(y_1), s_x(y_2), ..., s_x(y_{2^l})\}$) is set randomly, with equal probability and independently of selected k-bit master key $K^{(x)}$ value.

We use the properties of random homogeneous substitution for the analysis of round key sequences probability characteristics. For this purpose, on the set S'_n we define the uniform probabilistic distribution:

$$\forall i \in \{1,2,...,2^l\}, \forall u \in \{1,2,...,2^k\} : P(s_u(y_i)) = 2^{-l},$$

i.e. all probabilities of comparison of i-th block from y_i and u-th block from $s_u(y_i)$ are equal to each other and do not depend on i or $.u$.. Therefore, the probability of a random selection of substitution $s'_x \in S'_n \subset S_n$ (and the corresponding encryption master key) does not depend on the type of substitution, and defined as the inverse value of the power of the master keys set. Using (1) we get the next form:

$$P(s'_x) = \frac{P(s'_x(y_1))P(s'_x(y_2))\dots P(s'_x(y_{2^l}))}{\sum_{u=1}^{2^k} P(s'_u(y_1))P(s'_u(y_2))\dots P(s'_u(y_{2^l}))} = \frac{2^{-nl}}{\sum_{u=1}^{2^k} 2^{-nl}} = 2^{-k}. \quad (3)$$

Applying the considered model of random homogeneous substitution to analyse the probability properties of the key schedule elements we can estimate probabilities of coincidence of individual cyclic keys and their sequences, assuming that the round keys are generated randomly, equiprobable and independent from each other.

3 Probabilistic Properties of cycle keys

Let us introduce the following notations. Let randomly, equiprobably (probability is equal to 2^{-k}) and independently generated master key $K^{(x)}$ with length k bit be the input of key schedule construction (fig. 1). Then we note the sequence of t formed round keys as $K_{p\kappa}^{(x)} = \{K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}\}$, where every $K_i^{(x)}$ is the i-th cyclic key with length l-bit.

Let us assume that the cyclic keys $K_i^{(x)}$ are independent implementations of a random homogeneous substitution in the *i*-th point:

$$\forall i \in \{1,2,...,t\} : K_i^{(x)} = s'_x(y_i),$$

i.e. they are generated randomly, equiprobably and independently from each other, and for every $K_i^{(x)}$ the specific implementation of a random substitution $s'_x \in S'_n \subset S_n$ (implementation of the vector $\{s'_x(y_1), s'_x(y_2), ..., s'_x(y_n)\}$) is independent of $i \in \{1,2,...,t\}$.

Consider the probabilistic properties of randomly generated round key $K_i^{(x)}$ values for some fixed i. We estimate the average number of different values of cyclic key can be formed using all 2^k values of master keys $K^{(x)}$.

Lemma. The average number of different l-bit cyclic keys $K_i^{(x)}$ formed by 2^k implementations of the random homogeneous substitution is defined by expression:

$$N(k,l) = 2^l (1-(1-2^{-l})^{2^k}) \approx 2^l \left(1-\left(\frac{1}{e}\right)^{2^{k-l}}\right) \approx 2^l \left(1-(0,37)^{2^{k-l}}\right). \quad (4)$$

Proof. If the formation scheme of every cyclic key from the sequence $K_{p\kappa}^{(x)}$ is described by the model of a random homogeneous substitution with probability (1) of selection $s'_x \in S'_n \subset S_n$ (by the entered master key $K^{(x)}$), then, by definition, for any fixed $i \in \{1,2,...,t\}$ the probability of $K_i^{(x)}$ does not depend on $y_i \in Y = \{y_1, y_2, ..., y_{2^l}\}$ or $K^{(x)}$ and it is defined as the inverse of the power of the set Y, i.e. it is equal to $P(s'_x(y_i) = K_i^{(x)}) = 2^{-l}$. Master keys $K^{(x)}$ are selected independently from each other and corresponding events $s'_x(y_i) = K_i^{(x)}$ are independent. Therefore, we can use the formula for finding the probability of target event M times in N tests (Bernoulli formula):

$$P(N,M) = C_N^M (1-P(s'_x(y_i) = K_i^{(x)}))^{N-M} (P(s'_x(y_i) = K_i^{(x)}))^M = C_N^M (1-2^{-l})^{N-M} (2^{-l})^M.$$

The value $P(N,M)$ specifies the probability that at N independent implementations of random homogeneous substitution in i-th point a specific round key $K_i^{(x)} = s'_x(y_i)$ appears exactly M times. The value

$$P(2^k, 0) = (1-2^{-l})^{2^k}$$

specifies probability that at $N = 2^k$ independent implementations of random substitution in i-th point (in full set of master keys $K^{(x)}$ values) the round key $K_i^{(x)} = s'_x(y_i)$ will not appear a single time.

Inverse value

$$P(2^k, >0) = 1 - P(2^k, 0) = \sum_{i=1}^{2^k} C_{2^k}^i (1-2^{-l})^{2^k - i} (2^{-l})^i = 1-(1-2^{-l})^{2^k} \quad (5)$$

specifies probability of an event when at 2^k independent tests the round l- bit length key $K_i^{(x)}$ will be formed at least once.

Power of different *ℓ-bit* values set is equal to 2^ℓ where each of these values in 2^k independent implementations of the vector (2) appears at least once in i-th point of random substitution with prob-

ability (5). I.e. for 2^k different master keys $K^{(x)}$ defining vector (2) implementation by the key schedule construction it will be formed in average

$$N(k,l) = 2^l P(2^k, > 0) = 2^l (1-(1-2^{-l})^{2^k})$$

different round keys $K_i^{(x)}$. Using substitution $(1-2^{-l})^{2^l} \approx e^{-1}$ gives us a simplified formula in the right side of the expression (3), and, thus, completes the proof.

For the most simple case $k = l$ (equality of ciphertext block length to key length) the probability (5) gets the form

$$P(2^l, > 0) = \sum_{i=1}^{2^l} C_{2^l}^i (1-2^{-l})^{2^l - i} (2^{-l})^i = 1 - P(2^l, 0) = 1 - (1-2^{-l})^{2^k} \approx 1 - e^{-1} \approx 0,63$$

and the ratio of the average number $N(k,l)$ of different round keys $K_i^{(x)}$ to the number of 2^k different master keys $K^{(x)}$ under $k = l$ is determined as

$$\delta(k,l) = \frac{N(k,l)}{2^k} = \frac{2^l (1 - P(2^k, 0))}{2^k} = P(2^l, > 0) \approx 1 - e^{-1} \approx 0,63 \qquad (6)$$

what corresponds to the formula (27) in [15].

Under $k \neq l$ formula (6), as well as formula (27) in [15], is not satisfied, and we need to estimate the ratio $\delta(k,l)$ according to the general formula

$$\delta(k,l) = \frac{N(k,l)}{2^k} = 2^{l-k}(1-(1-2^{-l})^{2^k}) \approx 2^{l-k}\left(1 - \left(\frac{1}{e}\right)^{2^{k-l}}\right) \approx 2^{l-k}\left(1 - (0,37)^{2^{k-l}}\right). \quad (7).$$

Let us consider an example of using these relations.

Fig. 2 and 3 show dependency of the probabilities (5) and the relationships (7) for the blocks of length $0 \le l \le 16$ and keys $0 \le k \le 16$. It is obvious that even for such small lengths l and k which do not exceed 16 bits, there is a sharp transition from very small values (almost equal to zero values $P(2^k, > 0)$ and $\delta(k,l)$), to very large values (close to unity). This is true for the dependency $P(2^k, > 0)$, and the multiplier 2^{l-k} in (7) smoothes the final function (7), inverting the high-quality form of the dependency (5).

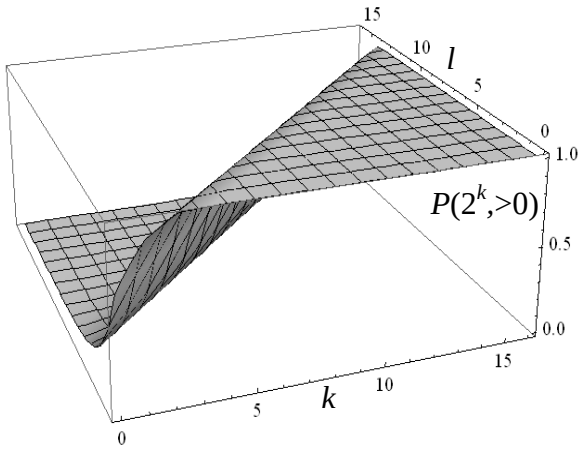

Fig. 2 - Dependence $P(2^k, > 0)$, if $0 \le l \le 16$ and $0 \le k \le 16$

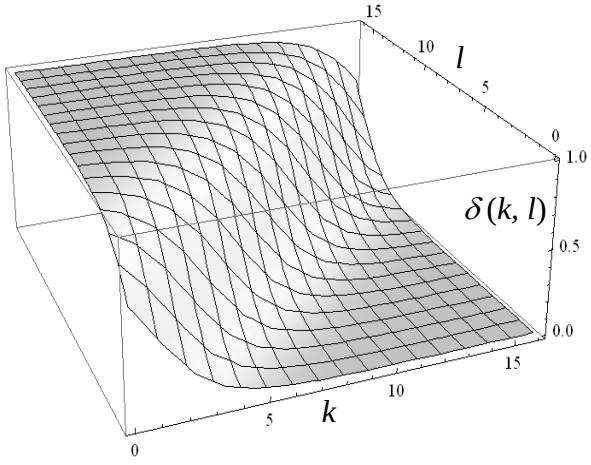

Fig. 3 - Dependence $\delta(k,l)$, if $0 \le l \le 16$ and $0 \le k \le 16$

Table 1 summarizes the values of the ratio of the different cyclic keys $K_i^{(x)}$ average number $N(k,l)$ to the number 2^k of different master keys $K^{(x)}$ for the most common values l and k in modern BSC and their scale models.

Data presented in Table 1 is calculated by the simplified formula on the right hand side of formula (7) using the Wolfram Alpha system computing algorithms [1,7]. These calculated values show efficiency of obtained analytical formulas for the round keys probability characteristics estimation.

Table 1 - The ratio of the average number of different l-bit cyclic keys $K_i^{(x)}$ formed by 2^k implementation of the random homogeneous substitution to the power of different master keys $K^{(x)}$ set

	k =16	***k =32***	***k =64***	***k =128***	***k =256***	***k =512***
***l* = 16**	0,63	$1{,}52\cdot10^{-5}$	$3{,}55\cdot10^{-15}$	$1{,}93\cdot10^{-34}$	$5{,}66\cdot10^{-73}$	$4{,}89\cdot10^{-150}$
***l* = 32**	$1-7{,}63\cdot10^{-6}$	0,63	$2{,}33\cdot10^{-10}$	$1{,}26\cdot10^{-29}$	$3{,}71\cdot10^{-68}$	$3{,}20\cdot10^{-145}$
***l* = 64**	$1-1{,}78\cdot10^{-15}$	$1-1{,}16\cdot10^{-10}$	0,63	$5{,}42\cdot10^{-20}$	$1{,}59\cdot10^{-58}$	$1{,}38\cdot10^{-135}$
***l* = 128**	$1-9{,}63\cdot10^{-35}$	$1-6{,}31\cdot10^{-30}$	$1-2{,}71\cdot10^{-20}$	0,63	$2{,}94\cdot10^{-39}$	$2{,}54\cdot10^{-116}$
***l* = 256**	$1-2{,}83\cdot10^{-73}$	$1-1{,}85\cdot10^{-68}$	$1-7{,}97\cdot10^{-59}$	$1-1{,}47\cdot10^{-39}$	0,63	$8{,}64\cdot10^{-78}$
***l* = 512**	$1-2{,}44\cdot10^{-150}$	$1-1{,}60\cdot10^{-145}$	$1-6{,}89\cdot10^{-136}$	$1-1{,}27\cdot10^{-116}$	$1-4{,}32\cdot10^{-78}$	0,63

To estimate probabilistic properties of cyclic keys $K_{p\kappa}^{(x)}$ sequences we summarize the positions of the lemma proved above for random, equiprobable and independent values $K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}$. Let us estimate the average number of different sequences $K_{p\kappa}^{(x)}$ that is generated using all 2^k values of master keys $K^{(x)}$. The following theorem is true.

Theorem. The average number of different cyclic keys sequences $K_{p\kappa}^{(x)} = \{K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}\}$, that is formed by 2^k independent implementations of random homogeneous substitution in i-th point, $i \in \{1,2,...,t\}$, is defined by:

$$N(k,l,t) = 2^l\left(1-(1-2^{-tl})^{2^k}\right) \approx 2^{tl}\left(1-\left(\frac{1}{e}\right)^{2^{k-tl}}\right) \approx 2^{tl}\left(1-(0{,}37)^{2^{k-tl}}\right). \qquad (8)$$

Proof is a generalization of the lemma`s results in the case of $K_{p\kappa}^{(x)} = \{K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}\}$ sequences formation. Indeed, in accordance with the assumption of $K_i^{(x)}$ cyclic keys generating by the independent implementations of random homogeneous substitution in i-th point, for each $i \in \{1,2,...,t\}$ the probability of $K_i^{(x)}$ does not depend on $y_i \in Y = \{y_1, y_2, ..., y_{2^l}\}$ or $K^{(x)}$. This probability is equal to $P(s'_x(y_i) = K_i^{(x)}) = 2^{-l}$. The joint probability of independent events is the product of the probabilities of these events, i. e.:

$$P(K_{p\kappa}^{(x)} = \{K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}\}) = \prod_{i=1}^{t} P(s'_x(y_i) = K_i^{(x)}) = 2^{-tl}.$$

Master keys $K^{(x)}$ are selected independently from each other and the corresponding events $K_{p\kappa}^{(x)} = \{K_1^{(x)}, K_2^{(x)}, ..., K_t^{(x)}\}$ are independent too. Therefore, using the Bernoulli formula just as in the lemma proof, we obtain the expression

$$P(N,M,t) = C_N^M (1-2^{-tl})^{N-M} (2^{-tl})^M,$$

which specifies the probability of the case that in N independent implementations the sequence $K_{pк}^{(x)}$ would appear exactly M times. The value

$$P(2^k,>0,t)=\sum_{i=1}^{2^k}C_{2^k}^{i}(1-2^{-tl})^{2^k-i}(2^{-tl})^{i}=1-P(2^k,0,t)=1-(1-2^{-tl})^{2^k} \qquad (9)$$

gives the probability of the case when in 2^k independent tests the specific sequence $K_{pк}^{(x)}$ is formed at least once.

The power of different t-sequences sets of l-bit values is equal to 2^{tl}, and each of these sequences with the probability (9) appear at the output of the round key schedule construction at least once. I. e. for 2^k different master keys $K^{(x)}$ that specify the implementations of random homogeneous substitution the cyclic key schedule construction it will be formed in average

$$N(k,l,t)=2^{tl}(1-P(2^k,0,t))$$

different $K_{pк}^{(x)}=\{K_1^{(x)},K_2^{(x)},...,K_t^{(x)}\}$ sequences, what under $(1-2^{-l})^{2^l}\approx e^{-1}$ simplification gives the target formula (8). This theorem allows us to obtain expression to estimate the ratio of different $K_{pк}^{(x)}$ sequences average number to the power of different $K^{(x)}$ master keys set:

$$\delta(k,l,t)=\frac{N(k,l,t)}{2^k}=2^{tl-k}(1-(1-2^{-tl})^{2^k})\approx 2^{tl-k}\left(1-\left(\frac{1}{e}\right)^{2^{k-tl}}\right)\approx 2^{tl-k}\left(1-(0,37)^{2^{k-tl}}\right) \qquad (10)$$

For convenience of $\delta(k,l,t)$ relations calculations we can write formula (10) in a different way. In the majority of practically important cases of block cipher (for instance, in estimating the properties of the BSC "Kalyna" key schedule) the master key length k is a multiple of the block length l, i.e. the ratio $k=ml$ is true, what after substitution in (10) it gives

$$\delta(ml,l,t)=2^{l(t-m)}(1-(1-2^{-tl})^{2^{ml}})\approx 2^{l(t-m)}\left(1-\left(\frac{1}{e}\right)^{2^{l(m-t)}}\right)\approx 2^{l(t-m)}\left(1-(0,37)^{2^{l(m-t)}}\right). \qquad (11)$$

Formula (11) shows that increasing of the multiplicity m is equivalent, in a probabilistic sense, to the corresponding decreasing of the sequence $K_{pк}^{(x)}=\{K_1^{(x)},K_2^{(x)},...,K_t^{(x)}\}$ length t. And conversely, the increasing of round keys sequence length t decreases the probability (9) as well master key length. A typical demonstration of this effect would be symmetry of function graphs relative to values l and k (Fig. 2,3). In this sense, the calculated values $\delta(ml,l,t)$ for the case $l\in\{16,32\}$ and $m,t\in\{1,2,4,8,16\}$ can be obtained from the data in Table 1 when selecting column with symbols ml and rows with symbols tl. As an example, table 2 shows the calculated values1 $\delta(ml,l,t)$ for $l=32$, which fully comply to the data presented in Table 1.

Table 2 - The ratio of the average number of different cyclic keys sequences to the power of set of different master keys length of $l=32$

	m = 1	**m = 2**	**m = 4**	**m = 8**	**m = 16**
t = 1	0,63	$2{,}33\cdot10^{-10}$	$1{,}26\cdot10^{-29}$	$3{,}71\cdot10^{-68}$	$3{,}20\cdot10^{-145}$
t = 2	$1-1{,}16\cdot10^{-10}$	0,63	$5{,}42\cdot10^{-20}$	$1{,}59\cdot10^{-58}$	$1{,}38\cdot10^{-135}$
t = 4	$1-6{,}31\cdot10^{-30}$	$1-2{,}71\cdot10^{-20}$	0,63	$2{,}94\cdot10^{-39}$	$2{,}54\cdot10^{-116}$
t = 8	$1-1{,}85\cdot10^{-68}$	$1-7{,}97\cdot10^{-59}$	$1-1{,}47\cdot10^{-39}$	0,63	$8{,}64\cdot10^{-78}$
t = 16	$1-1{,}60\cdot10^{-145}$	$1-6{,}89\cdot10^{-136}$	$1-1{,}27\cdot10^{-116}$	$1-4{,}32\cdot10^{-78}$	0,63

[1] *Values* $\delta(ml,l,t)$ *in tables 2-4 are calculated using simplified formula e* $\delta(ml,l,t)\approx 2^{l(t-m)}\left(1-e^{-2^{l(m-t)}}\right)$

The calculated values $\delta(ml,l,t)$ for cases $l \in \{64,128,256\}$, $m \in \{1,2,4,8\}$ and $t \in \{1,2,4,8,16\}$ are shown in Table 3.

The calculated values in Table 3 improve data on $\delta(k,l,t)$ estimation in [15]. The conclusion about virtually identical of the round keys sequences powers and encryption master keys in [15] is true. Data in the Table 3 clearly confirms this pattern. For all considered and practically significant relationships l and k, when $t > m$ is true, the ratio of the average number of different round keys sequences to the power of the different master keys set only slightly differs from unity. With further increasing of the round key sequence t length this difference rapidly decreases.

Table 3 - The ratio of the average number of different round keys sequences to the power of different master keys set

	***m* = 1**	***m* = 2**	***m* = 4**	***m* = 8**
***l* = 64**				
***t* = 1**	0,63	$1{,}52\cdot10^{-5}$	$5{,}66\cdot10^{-73}$	$4{,}89\cdot10^{-150}$
***t* = 2**	$1 - 7{,}62\cdot10^{-6}$	0,63	$3{,}71\cdot10^{-68}$	$3{,}20\cdot10^{-145}$
***t* = 4**	$1 - 1{,}78\cdot10^{-15}$	$1 - 1{,}16\cdot10^{-10}$	0,63	$8{,}64\cdot10^{-78}$
***t* = 8**	$1 - 2{,}44\cdot10^{-150}$	$1 - 1{,}60\cdot10^{-145}$	$1 - 1{,}27\cdot10^{-116}$	0,63
***t* = 16**	$1 - 5{,}13\cdot10^{-290}$	$1 - 9{,}46\cdot10^{-271}$	$1 - 3{,}22\cdot10^{-232}$	$1 - 3{,}73\cdot10^{-155}$
***l* = 128**				
***t* = 1**	0,63	$2{,}94\cdot10^{-39}$	$2{,}54\cdot10^{-116}$	$1{,}89\cdot10^{-270}$
***t* = 2**	$1 - 1{,}47\cdot10^{-39}$	0,63	$8{,}64\cdot10^{-78}$	$6{,}44\cdot10^{-232}$
***t* = 4**	$1 - 1{,}27\cdot10^{-116}$	$1 - 4{,}32\cdot10^{-78}$	0,63	$7{,}45\cdot10^{-155}$
***t* = 8**	$1 - 9{,}46\cdot10^{-271}$	$1 - 3{,}22\cdot10^{-232}$	$1 - 3{,}74\cdot10^{-155}$	0,63
***t* = 16**	$1 - 5{,}26\cdot10^{-579}$	$1 - 1{,}79\cdot10^{-540}$	$1 - 2{,}07\cdot10^{-463}$	$1 - 2{,}78\cdot10^{-309}$
***l* = 256**				
***t* = 1**	0,63	$8{,}64\cdot10^{-78}$	$6{,}44\cdot10^{-232}$	$3{,}58\cdot10^{-540}$
***t* = 2**	$1 - 4{,}32\cdot10^{-78}$	0,63	$7{,}45\cdot10^{-155}$	$4{,}15\cdot10^{-463}$
***t* = 4**	$1 - 3{,}22\cdot10^{-232}$	$1 - 3{,}73\cdot10^{-155}$	0,63	$5{,}56\cdot10^{-309}$
***t* = 8**	$1 - 1{,}79\cdot10^{-540}$	$1 - 2{,}08\cdot10^{-463}$	$1 - 2{,}78\cdot10^{-309}$	0,63
***t* = 16**	$1 - 5{,}54\cdot10^{-1157}$	$1 - 6{,}42\cdot10^{-1080}$	$1 - 8{,}61\cdot10^{-926}$	$1 - 1{,}55\cdot10^{-617}$
***l* = 512**				
***t* = 1**	0,63	$7{,}46\cdot10^{-155}$	$4{,}15\cdot10^{-463}$	$1{,}28\cdot10^{-1079}$
***t* = 2**	$1 - 3{,}73\cdot10^{-155}$	0,63	$5{,}56\cdot10^{-309}$	$1{,}72\cdot10^{-925}$
***t* = 4**	$1 - 2{,}08\cdot10^{-463}$	$1 - 2{,}78\cdot10^{-309}$	0,63	$3{,}09\cdot10^{-617}$
***t* = 8**	$1 - 6{,}42\cdot10^{-1080}$	$1 - 8{,}60\cdot10^{-926}$	$1 - 1{,}55\cdot10^{-617}$	0,63
***t* = 16**	≈ 1	≈ 1	≈ 1	$1 - 4{,}79\cdot10^{-1237}$

To confirm the adequacy and accuracy of the obtained results and our conclusions driven by these results the numerical experiment was executed. The experiment essence is counting the ratios of

the average number of different round keys sequences to power 2^k of the set of different master keys. To simulate the random substitution a simple function of random number generation, integrated into the environment of rapid applications development Embarcadero RAD Studio for Microsoft Windows from Embarcadero Technologies company, was used [18]. Each observation included estimation of sample mean (empirical average) of 100 model implementations. Each model implementation included calculation of the ratio of the average number of different round keys sequences to power 2^k of the set of different master keys.

In the experiment, we estimated both the sample means $\delta^*(ml,l,t)$ and sample variance D when the sample size of 100 elements. The results are summarized in Table 4. The last column of this table shows the accuracy values ε of the estimated characteristics for a given level of significance $\alpha = 0,05$.

Table 4 - Results of experimental researches and their comparison with theoretical calculations

	$\delta(ml,l,t)$	$\delta^*(ml,l,t)$	D	ε
***l* = 4, *m* = 1**				
***t* = 1**	0,632121	0,6453125	0,005716	0,014818
***t* = 2**	0,969391	0,969125	0,001758	0,008218
***t* = 3**	0,998049	0,9983125	0,000110	0,002056
***t* = 4**	0,999878	0,999875	0,000008	0,000554
***l* = 4, *m* = 2**				
***t* = 1**	0,062500	0,062500	0	0
***t* = 2**	0,632121	0,633074	0,000372	0,003780
***t* = 3**	0,969391	0,969675	0,000103	0,001989
***t* = 4**	0,998049	0,997934	0,000008	0,000554
***l* = 4, *m* = 3**				
***t* = 1**	0,003906	0,003906	0	0
***t* = 2**	0,062500	0,062500	0	0
***t* = 3**	0,632121	0,632152	0,000023	0,000940
***t* = 4**	0,969391	0,969492	0,000007	0,000519
***l* = 8, *m* = 1**				
***t* = 1**	0,632121	0,632836	0,000360	0,003719
***t* = 2**	0,998049	0,998051	0,000007	0,000519
***t* = 3**	0,999992	0,999961	0,000002	0,000277
***l* = 8, *m* = 2**				
***t* = 1**	0,003906	0,003906	0	0
***t* = 2**	0,632121	0,632176	0,000002	0,000277
***t* = 3**	0,998049	0,998063	$2,98 \cdot 10^{-8}$	$3,38 \cdot 10^{-5}$

As can be seen from the values in Table 4, results of experimental research fully confirm the validity of theoretical assumptions. In all cases the calculated values $\delta(ml,l,t)$ and obtained empirical

data $\delta^*(ml,l,t)$ differ on not more than ε (the absolute value of the error), and the probability with which the specified accuracy is achieved (the accuracy estimation) is 0,95. Since the accuracy characterizes the repeatability and stability of experiments [19], it can be argued that in 95% of the cases the value $\delta^*(ml,l,t)$ will differ from $\delta(ml,l,t)$ less than ε.

4 Conclusions and prospects for further researches

Our research of BSC round keys probabilistic properties have shown that even under random, equiprobable and independent formations the used key sequences can be the same, what inevitably reduces the power of implemented encryption-decryption mapping sets.

To describe the round keys schedule construction an abstract model of random substitution parameterized by encryption master key value was used. The obtained analytical relations allow us to estimate the probability properties of BSC cycle keys. In particular, the probability of multiple matching of round keys for a given number of the random homogeneous substitution implementations (a given number of master keys) is defined by Bernoulli formula. This ratio gives us an estimate of the probability of events when the specific round key will be generated at least once on the all set of master keys, i.e., it allows us to estimate the average number of different round keys on the output of formation scheme. The final result is also generalized on sequences of arbitrary length round keys, i.e., we can get numerical estimates of the probability properties of all BSC key schedule elements using the defined model.

Calculated values of ratios of the average number of different round keys sequences to power of different master keys sets provided in Tables 2,3, give an idea about ciphering of all admissible encryption-decryption mappings set. In particular, the given calculated values for the most important practical cases when the lengths of data blocks l and the keys $k = ml$ indicate that the number of rounds $t < m$, with probability close to unity, the specific round keys sequence will not be formed on the all set of master keys. This is equivalent to the fact, that average number of different round keys sequences will be negligible compared to the power of the of different master keys, i.e., the large number of mappings "plaintext - ciphertext" from the all set 2^k of maps would not be realized. And conversely, for the case $t > m$ the ratio of the average number of different round key sequences to the power of the master keys set almost does not differ from unity. With a further increase of t this difference decreases rapidly and it must be assumed that in such key schedule all valid mappings "plaintext - ciphertext" from a complete set 2^k of maps will be implemented. The conducted simulation modeling of the "ideal" key schedule construction allowed to obtain empirical estimates that coincide with theoretical calculations by formulas (9) - (11), what confirms the reliability and validity of research results. In particular, for all investigated cases the calculated values and obtained empirical data do not differ significantly (the relative error value < 3%), and the probability that the specified accuracy is achieved (the estimation accuracy) is 0,95. Therefore, we can argue that in 95% of the cases the calculated values and empirical data differ by less than error value.

It should be noted that the obtained analytical expressions and shown calculated values correspond to the hypothetical case of random, equiprobable and independent formation of round keys, i.e. to the "ideal" key schedule in the probabilistic aspect. The actual key schedule constructions are based on deterministic algorithms, parameterized by value of the encryption master key. Therefore, the obtained estimates on a random homogeneous substitution should be used as the upper limits for the probabilistic properties of round key sequences: the key schedule of real BSC can only approach in its characteristics to this "ideal" case and it does not improve the given calculated values.

The practical impact of these results lies in their immediate interpretation to estimate the probability properties of key schedule elements in the new national standard of BSC of Ukraine [12]. If the assumption that the cyclic keys of BSC "Kalyna" are independent implementations of random homogeneous substitution (i.e. it is formed random, equiprobability and independently of each other) is true, then the conclusion about virtually identical the powers of round keys sequences and encryption master keys is theoretically proven and the calculated values shown in Table 3 clearly confirm this regularity. Using in the new standard the "ideal" key schedules in addition to providing

resistance of cipher to the related keys attacks and to the attacks on implementation [15] allows to fully implement the key space of master keys and a corresponding set of maps "plaintext - ciphertext".

As perspective directions for further research one can mention the search or, at least, estimation the probability properties of subsets of the so-called BSC equivalent keys, when several different by value master keys lead to the formation of identical cyclic keys sequences, giving the identical bijective encryption-decryption mappings. In other words, the existence of the equivalent keys subsets reduces the power of the "plaintext - ciphertext" maps set, and the number of the key information that is numerically equal to the certainty measure of secret encryption parameters is also reduced. In addition, the existence of several master keys which are different by value, but equivalent by encryption function can be used by an attacker to implement cryptanalytic attacks, for instance, based on the substitution of protected information by false data in the case using of BSC in generation of message authentification code mode.

References

[1] Gorbenko I.D. Prykladna kryptologija. Teorija. Praktyka. Zastosuvannja: Pidruchnyk dlja vyshhyh navchal'nyh zakladiv / I.D. Gorbenko, Ju.I. Gorbenko. – Harkiv: Fort, 2013. – 880 s.
[2] Biryukov A. Slide Attacks / A.Biryukov, D.Wagner // Fast Software Encryption: 6th International Workshop, FSE'99 Rome, Italy, March 24–26, 1999 Proceedings. – Springer Berlin Heidelberg, 1999. – P. 245 – 259.
[3] Chalermpong Worawannotai, Isabelle Stanton A Tutorial on Slide Attacks [Electronic Resource]. – Way of access: http://www.eecs.berkeley.edu/~isabelle/slideattacks.pdf. – Title from the screen.
[4] Biryukov A. Advanced Slide Attacks / A. Biryukov, D. Wagner // Advances in Cryptology – EUROCRYPT 2000: International Conference on the Theory and Application of Cryptographic Techniques Bruges, Belgium, May 14-18, 2000 Proceedings. – Springer Berlin Heidelberg, 2000. – P. 589 – 606.
[5] Biham E. New types of cryptanalytic attacks using related keys / E.Biham // Springer-Verlag. – 1994. – № 4. – P. 229 – 246.
[6] Ciet M., Piret G., Quisquater J.-J. Related-Key and Slide Attacks: Analysis, Connections, and Improvements (Extended Abstract) [Electronic Resource]. – Way of access: http://citeseer.ist.psu.edu. – 2002. – Universite catholique de Louvain, Louvain-la-Neuve, Belgium. – Title from the screen.
[7] Biryukov A. Related-Key Cryptanalysis of the Full AES-192 and AES-256/ A. Biryukov, D. Khovratovich // Springer Berlin Heidelberg. – 2009. – P. 1 – 8.
[8] Daemen J. AES proposal: Rijndael / J. Daemen, V. Rijmen [Electronic Resource]. – Way of access: http://www.nist.gov/aes. – 1998. – Title from the screen.
[9] FIPS-197: Advanced Encryption Standard (AES) // National Institute of Standards and Technology. – 2001 [Electronic Resource]. – Way of access: http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf. – Title from the screen.
[10] Polozhennja pro porjadok zdijsnennja kryptografichnogo zahystu informacii' v Ukrai'ni, zatverdzhene Ukazom Prezydenta Ukrai'ny vid 22 travnja 1998 roku N 505/98.
[11] Rozrobka novogo blokovogo symetrychnogo shyfru: Zvit za pershyj etap NDR «Algorytm» (promizhnyj) / nauk. ker. I. D. Gorbenko; AT «IIT». – Kharkiv, 2014. – Tom 4. – 304 s.
[12] Informacijni tehnologii'. Kryptografichnyj zahyst infomacii'. Algorytm symetrychnogo blokovogo peretvorennja: DSTU 7624:2014. – K.: Minekonomrozvytku Ukrai'ny, 2015. – 238 s. – (Nacional'nyj standart Ukrai'ny).
[13] Sachkov V. N. Vvedenie v kombinatornye metody diskretnoi matematiki / V.N. Sachkov. – Moskva: Nauka, 1982. – 384 s.
[14] Sachkov V. N. Veroyatnostnye metody v kombinatornom analize / V.N.Sachkov. – Moskva: Nauka, 1978. – 287 s.
[15] Olijnykov R. V. Metody analizu i syntezu perspektyvnyh symetrychnyh kryptografichnyh peretvoren': avtoref. dys. na zdobuttja nauk.stupenja d-ra tehn. nauk: 05.13.05 / R. V. Olijnykov; HNURE. – Harkiv, 2014. – 42 c. – ukr.
[16] NIST Special Publication 800-38D. Block Cipher Modes [Electronic Resource]. – Way of access: http://csrc.nist.gov/groups/ST/toolkit/BCM/ current_modes.html. – Title from the screen.
[17] Voprosno-otvetnaya sistema WolframAlpha Modes [Electronic Resource]. – Way of access: http://www.wolframalpha.com /. – Title from the screen.
[18] Integrirovannaya sreda razrabotki Embarcadero RAD Studio [Electronic Resource]. – Way of access: http://www.embarcadero.com/ products/rad-studio. – Title from the screen.
[19] Venttsel' E.S. Teoriya veroyatnostei i ee inzhenernye prilozheniya: Ucheb. posobie dlya vtuzov / E.S.Venttsel', L.A. Ovcharov.– 2-e izd., ster. – Moskva: Vyssh. shkola, 2000. – 480 s.

Надійшло: червень 2016.

Автори: Олександр Кузнецов, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет імені В.Н. Каразіна, Харків, Україна. E-mail: kuznetsov@karazin.ua
Юрій Горбенко, к.т.н., с.н.с., Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: YuGorbenko@iit.kharkov.ua
Євгенія Колованова, к.т.н., ст. викладач, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: e.kolovanova@gmail.com

Ключовий розклад блокових симетричних шифрів.

Анотація: Досліджуються комбінаторні властивості ключового розкладу блокових симетричних шифрів в припущенні, що циклові (раундові) ключі формуються випадково, рівноймовірно і незалежно один від одного. Для абстрактного опису такого формування використовується модель випадкової однорідної підстановки. Отримані аналітичні вирази дозволяють оцінити потужність множини реалізованих відображень зашифрування-розшифрування, отримати оцінки імовірнісних властивостей послідовностей раундових ключів і відносин середнього числа різних ключових послідовностей до потужності множини різноманітних майстер-ключів. Результати імітаційного моделювання підтверджують достовірність і обгрунтованість отриманих аналітичних виразів.

Ключові слова: ключовий розклад, циклові ключі, комбінаторні властивості, блокові симетричні шифри.

Рецензент: Виктор Долгов, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина.
E-mail: dolgovvi@mail.ru

Поступила: июнь 2016.

Авторы: Александр Кузнецов, д.т.н., проф., академик Академии наук прикладной радиоэлектроники, Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина. E-mail: kuznetsov@karazin.ua
Юрий Горбенко, к.т.н., с.н.с, Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина.
E-mail: YuGorbenko@iit.kharkov.ua
Евгения Колованова, к.т.н., ст. преподаватель, Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина. E-mail: e.kolovanova@gmail.com

Ключевое расписание блочных симметричных шифров.

Аннотация: Исследуются комбинаторные свойства ключевого расписания блочных симметричных шифров в предположении, что цикловые (раундовые) ключи формируются случайно, равновероятно и независимо друг от друга. Для абстрактного описания такого формирования используется модель случайной однородной подстановки. Полученные аналитические выражения позволяют оценить мощность множества реализуемых отображений зашифрования-расшифрования, получить оценки вероятностных свойств последовательностей раундовых ключей и отношений среднего числа различных ключевых последовательностей к мощности множества различных мастер-ключей. Результаты имитационного моделирования подтверждают достоверность и обоснованность полученных аналитических выражений.

Ключевые слова: ключевое расписание, цикловые ключи, комбинаторные свойства, блочные симметричные шифры.