# Entropically secure cipher for messages generated by Markov chains with unknown statistics

Boris Ryabko[1,2]

[1]Federal Research Center for Information and Computational Technologies , Novosibirsk, Russia
[2]Novosibirsk State University, Russia
boris@ryabko.net

**Abstract**

In 2002, Russell and Wang proposed a definition of entropic security, which was developed within the framework of secret-key cryptography. An entropically secure system is unconditionally secure, that is, unbreakable regardless of the adversary's computing power. The notion of an entropically secure symmetric encryption scheme is important for cryptography because one can construct entropically secure symmetric encryption schemes with keys much shorter than the length of the input, thus circumventing Shannon's famous lower bound on key length. In this report we suggest an entropically secure scheme for the case where the encrypted message is generated by a Markov chain with unknown statistics. The length of the required secret key is proportional to the logarithm of the message length (as opposed to the length of the message itself for the one-time pad).

**Keywords:** Information Theory, entropy security, indistinguishability, symmetric encryption scheme, unconditionally secure, Markov chain, unknown statistics.

## 1 Introduction

In 1949, K. Shannon, in his remarkable article [1], described the perfect secret system and showed that the one-time pad is such a system. Since then, it has been generally accepted that the length of the secret key should be equal to the length of the encrypted message (or at least its entropy). Russell and Wang [2] proposed the notion of entropic security, which gives a possibility to build a symmetric encryption scheme with a secret key much shorter than the length of the input, thus, in a sense, circumventing the mentioned Shannon's lower bound on key length. Informally, the entropy-secure symmetric encryption scheme uses the entropy of the input message to make the required secret key shorter.

The concept of entropic security has been generalized and developed by Dodis and Smith [3] and investigated by several other authors [4, 5, 6]. In order to describe it, suppose that there is a sender Alice and a receiver Bob who share a secret key $K$, and Alice wants to securely send some message $M$ to Bob over a public channel. The message $M$ is assumed to come from some a-priori distribution on $\Lambda^n$ where $\Lambda$ is a finite alphabet, $n \geq 1$, and $K$ is a sequence of equally probable and independent binary digits. Informally, the goal is to compute $E(M, K)$ which allows Bob to extract $M$ from $E(M, K)$ using $K$ and (the decoder) $D(E, K), (D(E, K) = M)$, in such a way as to reveal "no information" about $M$ to the adversary Eve beyond what she already knew. It is assumed that $E(M, K)$ is a probabilistic map, that is, it can also use random numbers, which are unknown to Bob.

The following formal definition of the entropic security belongs to Russell and Wang [2] (see also Dodis and Smith [3]):

**Definition 1.** *A probabilistic map $E(M, K)$ is said to hide all functions $f$ on $\Lambda^n$ to $\{0, 1\}^*$ with leakage $\epsilon, \epsilon > 0$, if, for every adversary $A$, there exists some adversary $\hat{A}$ (who does not know $E(M, K)$) such that for all functions $f$ from $\Lambda^n$ to $\{0, 1\}^*$,*

$$| Pr\{A(E(M, K)) = f(M)\} - Pr\{\hat{A}(\,) = f(M)\} | \leq \epsilon. \qquad (1)$$

*(Note that $\hat{A}$ does not know $E(M, K)$ and, in fact, she guesses the meaning of the function $f(M)$, ignoring $E(M, K)$.)*

*The cipher $E(M, K)$ is $\epsilon$-entropically secure for a probability distribution $P$ on $\Lambda^n$ if $E(M, K)$ hides all functions $f$ on $\Lambda^n$ to $\{0, 1\}^*$ with leakage $\epsilon$ when $M$ obeys the distribution $P$.*

Another concept, namely, that of indistinguishability, provides another way evaluate the strength of the cipher. To describe it, we first need to define min-entropy.

For a probability distribution $P$ on the alphabet $S$ the min-entropy is defined as follows:

$$h_{min}(P) = -\log \max_{a \in S} P(a)\,, \qquad (2)$$

$\log = \log_2$.

**Definition 2.** *(Dodis and Smith [3].) A randomized map $Y()$ is $(t, \epsilon)$-indistinguishable if there is a random variable $G$ such that for every distribution on a set $\mathbf{M}$ with min-entropy at least $t$, we have*

$$SD(Y(M), G) \leq \epsilon,$$

*where for two probability distributions* $A, B$

$$SD(A,B) = \frac{1}{2} \sum_{M \in \mathbf{M}} |Pr\{A = M\} - Pr\{B = M\}| .$$

Informally, in what follows the map $Y(\,)$ will be the cipher and, again, $G$ does not depend on the ciphered message. So, Eve can guess the message regardless of its cipher.

Dodis and Smith [3] showed that entropy security and indistinguishability are equal (up to small constants in key length). In particular, they show that if a cipher is $\epsilon$-entropically secure, it is $4\epsilon$-indistinguishable.

The main result of this paper is as follows: We describe an $\epsilon$-entropically secure cipher for the case where the probability distribution $\mu$ is unknown, but it is known that it belongs to class of stationary ergodic Markov chains with finite memory, or connectivity, $m$, $m \geq 0$, whose definition is given in Appendix. (If $m = 0$ then the symbols generated by $\mu$ are independent and identically distributed – i.i.d.). The length of the required secret key is $c_1 \log n + c_2 \log(1/\epsilon) + c_3$, where $n$ is the length of encrypted sequence, $c_1, c_2$ and $c_3$ are constants that depend on $m$ and the size of the alphabet $\Lambda$. (Recall that all participants know $m$, but the secret key are known only to Alice and Bob and the key is used only once).

The proposed method is based on the concept of the $\epsilon$-entropically secure cipher and some results of universal coding, which makes it possible to efficiently "compress" messages with unknown statistics [7].

## 2 Preliminaries

### 2.1 Universal coding

First, we consider the simplest case where the alphabet is $\{0,1\}^n, n \geq 1$ and letters are generated by some i.i.d. source $\mu$ and $\mu(0)$, $\mu(1)$ are unknown. The goal is to build a lossless code which "compresses" $n$-letter sequences in such a way that the average length (per letter) of the compressed sequence is close to the Shannon entropy $h(\mu)$, which is the lower limit of the code-word length (lossless code is such that the encoded messages can be decoded without errors and $h(\mu) = -(\mu(0)\log\mu(0) + (1-\mu(0))\log(1-\mu(0))\ )$ [7, 8].

The first universal code was invented by Fitingoff [9] and we use this code as a part of the suggested entropically secure cipher. In order to describe this code we consider any word $v \in \{0,1\}^n$ and denote by $\nu$ the number of ones in $v$ and let $S_\nu$ be the set of n-length words with $\nu$ ones. Fitingoff proposed

to encode the word $v$ by two subwords $u$ (prefix) and $w$ (suffix), where $u$ is the binary notation of an integer $\nu$ and $w$ is the index of the word $v$ in the subset $S_\nu$. It is assumed that the words in $S_\nu$ are ordered 0 to $(|S_\nu| - 1)$ (say, lexicographically) and the lengths of $u$ and $w$ are equal to $\lceil \log(n+1) \rceil$ and $\lceil \log |S_n| \rceil$, respectively. For example, for $n = 3$, $v = 100$ we obtain $\nu = 1, u = 01, w = 10$.

Recall the definition of the so-called prefix-free code. A set of words $U$ is prefix-free if for any $u, v \in U$ neither $u$ is a prefix of $v$ nor $v$ is a prefix of $u$ [8]. Clearly, the Fitingoff code is prefix-free. If some code $\lambda$ is prefix-free, then for any sequence $x_1x_2....x_n, n \geq 1$, $x_i \in \Lambda$, the encoded sequence $\lambda(x_1)\lambda(x_2)...\lambda(x_n)$ can be decoded to $x_1x_2....x_n$ without errors. Hence, any prefix-free code is a lossless one.

If we denote the Fitingoff code by $code_F$ we obtain from its description

$$|code_F(v)| = \lceil \log(n+1) \rceil + \lceil \log |S_\nu| \rceil \ + 1 \, . \tag{3}$$

For this code the ability to compress messages is based on the simple observation that probabilities of all messages from $S_\nu$ are equal for any distribution $\mu$ and, hence, $\mu(v) \leq 1/|S_\nu|$ for $\mu$ and any word $v \in S_\nu$. From this inequality and (3) we obtain

$$|code_F(v)| \leq \log(n+1) + 3 + \log(1/\mu(v)) \, . \tag{4}$$

(Let's explain the name "universal code." Clearly, the average code-length $E_\mu(|code_F|)$ is not grater than $\log(n+1)+3+nh(\mu)$ and, hence, the average length per letter $E_\mu(|code_F|)/n$ is not grater than $h(\mu) + (\log n + 3)/n)$. We can see that $E_\mu(|code_F|)/n \to h(\mu)$ if $n \to \infty$. So, one code compresses sequences generated by any $\mu$, that is, the code universal.)

The Fitingoff code described generalizes to i.i.d. processes with any finite alphabet $\Lambda$, as well as to Markov chains with memory or connectivity $m$, based on the same method as for binary i.i.d. [7]. Namely, the set of all $n$-letter words is divided into subsets of equiprobable words, and the code of any word is represented by a prefix and a suffix, where the prefix contains the number of the set with equiprobable words which contains the encoded one, and the prefix is the number in this set. It can be shown that the number of sets with equiprobable words is bounded above by $(|\Lambda| - 1)|\Lambda|^m$ ([7, 8]), and similarly (4) we can deduce that

$$|code_F(v)| \leq \log((|\Lambda| - 1)|\Lambda|^m) + 3 + \log(1/\mu(v)) \, . \tag{5}$$

It is important to note that there exists an algorithm to find the codewords which is based on method of fast calculation of numbers in $S_\nu$, see [10]. The complexity of this algorithm is $O(n \log^3 n \log \log n)$.

### 2.2 Entropically secure ciphers

Dodis and Smith [3], based on the results of Russell and Wang [2], proved the following

**Theorem (Russell-Wang, Dodis- Smith )** ([2], [3]). Let there be a probability distribution $\sigma$ on an alphabet $\Lambda = \{0,1\}^l$, $l \geq 1$. Then, for any $\epsilon > 0$, there exists an $\epsilon$- entropically secure cipher $E(M,K)$, $M \in \{0,1\}^l$ with the length of the key

$$|K| = l - h_{min}(\sigma) + 2log(1/\epsilon) + 2. \tag{6}$$

Take any such cipher and denote it $cipher_{RW-DS}(M,K)$. Dodis and Smith described three algorithm of such ciphers with a key length (6) whose complexity grows polynomially in $l$ and $\log(1/\epsilon)$ (One such a cipher is described in Appendix).

It is important to note that each of the three constructions of the ciphers depends only on min-entropy, that is, the cipher construction is the same for all distributions with the same min-entropy (but, of course, depends on $\epsilon$ and $l$).

## 3 The cipher

### 3.1 Randomised prefix-free codes

Let $\lambda$ be a prefix-free code for some alphabet $\Lambda^*$ and $L = \max_{a\in\Lambda^*} |\lambda(a)|$. The randomized code $\rho_\lambda$ maps elements from $\Lambda^*$ to the set $\{0,1\}^L$ defined as follows:

$$\rho_\lambda(a_i) = \lambda(a_i)\, r^i_{|\lambda(a_i)|+1} r^i_{|\lambda(a_i)|+2} ... r^i_L\,, \tag{7}$$

where $r^i_{|\lambda(a_i)|+1}, r^i_{|\lambda(a_i)|+2}, ..., r^i_L$ are uniformly distributed and independent random bits (for all $i$).

Let us define the probability distribution $\pi_{\lambda,\mu}$ on $\{0,1\}^L$ as follows:

$$\pi_{\lambda,\mu}(y_1 y_2 ... y_L) = \mu(a) 2^{-(L-|\lambda(a)|)}$$

$$\text{if} \quad y_1 y_2 ... y_{|\lambda(a_i)|} = \lambda(a). \tag{8}$$

If for some $y = y_1 ... y_L$ any $\lambda(a)$ is not a prefix of $y$, then $\pi_{\lambda,\mu}(y) = 0$.

Let us estimate the min-entropy of the distribution $\pi_{\lambda,\mu}$. From this equation and the definition of the min-entropy (2) we obtain the following:

$$h_{min}(\pi_{\lambda,\mu}) = L - \max_{a\in\Lambda}(|\lambda(a)| - \log(1/\mu(a))\,. \tag{9}$$

Now we consider the Fitingoff code applied to $n$-letter sequences generated by a Markov chain $\mu$ of memory $m$ over some alphabet $\Lambda$. The Fitigoff code is prefix-free and, hence, from (5) and (9) we obtain the following

**Statement.** For any distribution $\mu$

$$h_{min}(\pi_{code_F,\mu}) > L - (|\Lambda|^m(|\Lambda| - 1)\log n \, + \, 3)\,. \tag{10}$$

In particular, for an i.i.d. source with binary alphabet

$$h_{min}(\pi_{code_F}) > L - (\log n + 3)\,.$$

### 3.2 Description of the cipher

Here we describe a cipher with the key of length $const_1 \log n + const_2 \log(1/\epsilon) + const_3$, which is $\epsilon$-entropically secure for $n$-letter sequences generated by any (unknown) Markov chain $\mu$ of memory $m$ over some alphabet $\Lambda$.

Briefly, the encryption is done as follows: first compress the message with the Fitingoff code, then randomize the encoded message according to (7) and then encrypt the received $\rho_{code_F,\mu}(\,)$ with an entropically secure cipher. (Note that the distribution of $\mu$ is unknown.)

In detail, this algorithm is as follows:

**Parameters:** $\epsilon > 0$, the alphabet $\Lambda$, the memory of Markov chain $m$ and the length of the ciphered message $n$.

**Input:** a word $v \in \Lambda^n$.

**1st step:** Encode $v$ with the Fitigoff code $code_F(v)$ (with parameters $\Lambda, m$ and $n$).

**2nd step:** Calculate the random word $\rho_{code_F}(v)$ ($\in \{0,1\}^L$).

**3rd step:** Calculate the $\epsilon$-entropically secure cipher $cipher_{RW-DS}(\rho_{code_F}(v), K)$ with the length of the secret key $|K| = (|\Lambda|^m(|\Lambda| - 1)\log n + 2\log(1/\epsilon) + 5$ bits.

**Output:** $cipher_{RW-DS}(\rho_{code_F}(v))$.

The decryption algorithm is as follows: first Bob decrypts the word $E(\rho_{code_F}(v), K)$ $(= cipher_{RW-DS}(\rho_{code_F}(v))\,)$ with the known secret key $K$ and obtains the word $\rho_{code_F}(v)$. Then, based on the prefix-free property of the Fitingoff code, Bob finds the word $code_F(v)$ and then decodes it to get $v$.

The described cipher uses compression and randomisation. Denote it $cipher_{c\&r}$.

The theorem of Russell-Wang and Dodis-Smith guarantees the entropic security and indistinguishability for the first cipher $cipher_{RW-DS}$, so, we

need to prove a similar property for the proposed $cipher_{c\&r}$. Despite the equivalence of the concepts of entropic security and indistinguishability [3], we will prove these properties separately due to the great importance of this fact for the described cipher $cipher_{c\&r}$.

The following theorem describes the entropic security property for this cipher:

**Theorem 1.** *Let $\epsilon > 0$ and suppose that the cipher $cipher_{c\&r}$ is applied to $n$-letter words $M$ generated by a stationary ergodic Markov chain with memory $m, m \geq 0$, and an alphabet $\Lambda$, and let the length of the secret key $K$ be $(|\Lambda|^m(|\Lambda| - 1)\log n + 2\log(1/\epsilon) + 5$. Then $cipher_{c\&r}$ is $\epsilon$-entropically secure, that is, for any function $A : \{0,1\}^L \to \{0,1\}^*$ and $f : \Lambda^n \to \{0,1\}^*$ there exists such a function $\hat{A} : \{0,1\}^L \to \{0,1\}^*$ that*

$$|Pr\{A(cipher_{c\&r}(M,K) = f(M)\} - Pr\{\hat{A}(\,) = f(M)\}| \leq \epsilon,$$

*where $\hat{A}$ does not use $cipher_{c\&r}(M)$.*

*Proof.* The cipher $cipher_{RW-DS}(\rho_{code_F}(v), K)$ with the length of the secret key $|K| = (|\Lambda|^m(|\Lambda| - 1)\log n + 2\log(1/\epsilon) + 5$ is applied to $\{0,1\}^L$ (see the step 3). First we note that the cipher is $\epsilon$-entropically secure. Indeed, from Theorem of Russell-Wang and Dodis- Smith (see (6)) and the estimate of the min-entropy (10) we can see that such a cipher exists for the distribution $\pi_{code_F,\mu}$ for any (unknown) $\mu$. So, from the definition of $\epsilon$-entropical security we can see that for any function $g$

$$|Pr\{A(cipher_{RW-DS}(v) = g(v)\} - Pr\{\hat{A}(\,) = g(v)\}| \leq \epsilon,$$

where $v, v \in \{0,1\}^L$, $g$ is any function defined on $\{0,1\}^L$ ($g : \{0,1\}^L \to \{0,1\}^*$) and $\hat{A}(\,)$ does not depend on $v$ (to be short, $\lambda = code_F$). Taking into account that the code $\lambda$ is prefix-free, we can define such a function $\phi$ that for any $a \in \Lambda^n$ and $u = \rho_\lambda(a)$, $\phi(u) = a$. For any function $f : \Lambda^n \to \{0,1\}^*$ and $M$ consider the function $g(\rho_\lambda(M)) = f(\phi(\rho_\lambda(M))(= f(M))$. This equation is valid for the function $g$ and for $v = \rho_\lambda(M)$, hence

$$|Pr\{A(cipher_{ds}(\rho_\lambda(M)) = f(\phi(\rho_\lambda(M))\}-$$

$$Pr\{\hat{A}(\,) = f(\phi(\rho_\lambda(M))\}| \leq \epsilon.$$

Taking into account that $cipher_{c\&r}(M) = cipher_{RW-DS}(\rho_\lambda(M))$ and $f(\phi(\rho_\lambda(M)) = f(M)$, we can see from the latter inequality that

$$|Pr\{A(cipher_{c\&r}(M)) = f(M)\}-$$

$$Pr\{\hat{A}(\,) = f(M)\}| \leq \epsilon\,.$$

The theorem is proven.

The following theorem establishes indistinguishability of $cipher_{c\&r}$.

**Theorem 2.** *Let $\epsilon > 0$ and and suppose that the cipher $cipher_{c\&r}$ is applied to $n$-letter words $M$ generated by a stationary ergodic Markov chain with memory $m, m \geq 0$, and an alphabet $\Lambda$, and let the length of the secret key $K$ be $(|\Lambda|^m(|\Lambda| - 1)\log n + 2\log(1/\epsilon) + 5$. Then, this cipher is $4\epsilon$-indistinguishable.*

*Proof.* The cipher $cipher_{RW-DS}$ is $\epsilon$-entropically secure (see Theorem 1). As we mentioned in Introduction, Dodis and Smith [3] showed that it means that this cipher is $4\epsilon$-indistinguishable. Our goal is to prove this property for $cipher_{c\&r}$. The $4\epsilon$-indistinguishability means that $SD(cipher_{RW-DS}, G) \leq 4\epsilon$, where $G$ is a random variable on $\{0,1\}^L$ (which is independent on $cipher_{RW-DS}$).

Define $U_a = \{cipher_{RW-DS}(\lambda(a)\, r) : r \in \{0,1\}^{L-\lambda(a)}\}$ and let the a random variable of $G'(v)$ be defined as follows:

$$Pr\{G' = v\} = \sum_{w \in U_v} Pr\{G = w\}.$$

The following chain of equalities and inequalities is based on these definitions and the triangle inequality for $L_1$:

$$SD(cipher_{c\&r}, G') =$$

$$\frac{1}{2}\sum_{u \in \Lambda^n} |Pr\{cipher_{c\&r} = u\} - Pr\{G' = u\}| =$$

$$\frac{1}{2}\sum_{v \in \{0,1\}^n} |\sum_{w \in U_v} (Pr\{cipher_{RW-DS} = w\} - Pr\{G = w\})| \leq$$

$$\frac{1}{2}\sum_{v \in \Lambda^n}\sum_{w \in U_v} |Pr\{cipher_{RW-DS} = w\} - Pr\{G = w\}\,| =$$

$$\frac{1}{2}\sum_{w \in \{0,1\}^L} Pr\{cipher_{RW-DS} = w\} - Pr\{G = w\}\,| =$$

$$SD(cipher_{RW-DS}, G) \leq 4\epsilon\,.$$

So, $SD(cipher_{c\&r}, G') \leq 4\epsilon$.

Theorem is proven.

Let us estimate the complexity of encoding and decoding. As we mentioned above, the encoding and decoding fitting complexity is $O(n \log^{const})$. The complexity of the Dodis and Smith cipher is polynomial in $n$. Thus, the complexity of the proposed cipher is also polynomial in $n$.

# 4 Appendix

## 4.1 The definition of a stationary ergodic Markov chain with memory, or connection, $m$.

First we give a definition of stationary ergodic processes. The time shift $T$ on $\Lambda^\infty$ is defined as $T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$. A process $P$ is called stationary if it is $T$-invariant: $P(T^{-1}B) = P(B)$ for every Borel set $B \subset \Lambda^\infty$. A stationary process is called ergodic if every $T$-invariant set has probability 0 or 1: $P(B) = 0$ or 1 whenever $T^{-1}B = B$ [11, 12].

We denote by $M_\infty(\Lambda)$ the set of all stationary and ergodic sources and let $M_0(\Lambda) \subset M_\infty(\Lambda)$ be the set of all i.i.d. processes. We denote by $M_m(\Lambda) \subset M_\infty(\Lambda)$ the set of Markov sources of order (or with memory, or connectivity) not larger than $m$, $m \geq 0$. By definition $\mu \in M_m(\Lambda)$ if

$$\mu(x_{t+1} = a_{i_1} | x_t = a_{i_2}, x_{t-1} = a_{i_3}, \dots, x_{t-m+1} = a_{i_{m+1}}, \dots)$$

$$= \mu(x_{t+1} = a_{i_1} | x_t = a_{i_2}, x_{t-1} = a_{i_3}, \dots, x_{t-m+1} = a_{i_{m+1}})$$

for all $t \geq m$ and $a_{i_1}, a_{i_2}, \dots \in \Lambda$.

## 4.2 Entropically secure ciphers.

In this part we describe one entropically secure cipher from [3], part 3.2.

Let $\{h_i\}_{i \in I}$ be some family of functions $h_i : \{0,1\}^k \to \{0,1\}^n$, indexed over the set $I = \{0,1\}^r$. By definition, a collection of functions from $n$-bit words to $n$-bits is XOR-universal if:

$$\forall a, x, y \in \{0,1\}^n, x \neq y, Pr\{h_i(x) \oplus h_i(y) = a\} \leq \frac{1}{2^{n-1}},$$

if $i$ is randomly chosen from $I$ according to the uniform distribution ($\oplus$ is symbol-by-symbol modulo 2 summation). Also, suppose that there is a XOR-universal collection of functions whose description is public and, hence, it is known to Alice, Bob and Eve.

Dodis and Smith consider an encryption scheme of the form

$$E(m, K, i) = (i; m \oplus h_i(K)$$

where $i$ is randomly chosen from $I$ according to the uniform distribution, and $K$ is a $k$-bit secrete key. Note that $m$ is a ciphered message of length $n$, $i$ is the number of $h_i$ in the set $I$ and $i = \log |I| = r$. (Dodis and Smith notice that this scheme is a special low-entropy, probabilistic one-time pad.) Decryption is obviously possible, since the description of the function $h_i$ is public. It is shown [3] that this cipher is $\epsilon$-entropically secure for $|k| \geq n - h_{min} + 2\log(1/\epsilon) + 2$ if the function family $\{h_i\}_{i \in I}$ is XOR-universal.

An example of XOR-universal family is as follows [3]: View $\{0,1\}^n$ as $\mathcal{F} = GF(2^n)$, and embed the key set $\{0,1\}^k$ as a subset of $\mathcal{F}$. For any $i \in \mathcal{F}$, let $h_i(K) = iK$, with multiplication in $\mathcal{F}$. This yields a family of linear maps $\{h_i\}$ with $2^n$ members. For this family the complexity of ciphering and deciphering is $O(n \log n \log\log n)$ [3].

It is important to note that the length of the secret key ($k$) depends only on the min-entropy of the probability distribution and does not depend on other parameters of the distribution.

# References


[1] Shannon C. E., "Communication theory of secrecy systems", *The Bell system technical journal*, 1949, 656-715.

[2] Russell A, Wang H., "How to fool an unbounded adversary with a short key", *IEEE Transactions on Information Theory*, 2006, 1130-1140.

[3] Dodis Y., Smith A., "Entropic security and the encryption of high entropy messages", Theory of Cryptography Conference, 2005, 556-577.

[4] Li X., Tang Q., Zhang Z., "Fooling an Unbounded Adversary with a Short Key, Repeatedly: The Honey Encryption Perspective", 2nd Conference on Information-Theoretic Cryptography (ITC 2021), 2021. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 1-21.

[5] Juels A, Ristenpart T., "Honey encryption: Security beyond the brute-force bound", Annual international conference on the theory and applications of cryptographic techniques, 2014 May 11, 293-310.

[6] Ryabko B, "Using data compression and randomization to build an unconditionally secure short key cipher", Cryptology ePrint Archive, Report 2021/1667, 2021.

[7] Krichevsky R., *Universal compression and retrieval*, Springer Science & Business Media, 1994.

[8] Cover T. M. , Thomas J. A., *Elements of information theory*, Wiley-Interscience, 2006..

[9] Fitingof B. M., "Optimal coding in the case of unknown and changing message statistics", *Problemy Peredachi Informatsii, 2(2)*, 1966, 3-11.

[10] Ryabko B.Ya., "The fast enumeration of combinatorial objects", *Discrete Math.and Applications, v.10, n. 2*, 1998, 163-182.

[11] Billingsley P., *Ergodic Theory and Information*, John Wiley and Sons, Inc., 1965.

[12] Ryabko D., *Asymptotic nonparametric statistical analysis of stationary time series*, Springer, 2019.