LOWER BOUNDS FOR THE NUMBER OF KEYS IN A ZAKREVS'KIJ CIPHER

D. S. Ananichev* and D. D. Dublemenkhy*

UDC 519.175.3

We present a new lower bound for the number of pairwise nonequivalent invertible Mealy machines with strongly connected underlying digraphs. Bibliography: 2 titles.

1. Notation

A cryptomachine $C = (X, Q, Y, K, \psi, \varphi)$ is defined by specifying a finite set $Q$ of states, finite alphabets $X$ and $Y$, a finite set $K = Q \times K_0$ of keys, a transition function $\psi : Q \times X \times K_0 \rightarrow Q$, and an output function $\varphi : Q \times X \times K_0 \rightarrow Y$.

If we fix a key $(q_0, k) \in Q \times K_0$, then we obtain the Mealy machine $C(q_0, k) = (Q, X, Y, \psi, \varphi, q_0)$, where the transition function $\psi_k : Q \times X \rightarrow Q$ is defined as $\psi_k(q, x) = \psi(q, x, k)$ for any $q \in Q$, $x \in X$, the output function $\varphi_k : Q \times X \rightarrow Y$ is defined as $\varphi_k(q, x) = \varphi(q, x, k)$ for any $q \in Q$, $x \in X$, and $q_0$ is the initial state. So $K_0$ is a kind of set of Mealy machines without a defined initial state. For every $q \in Q$, the function $\varphi_k : X \rightarrow Y$ is defined as $\varphi_k(x) = \varphi_k(q, x)$ for every $x \in X$. We say that a Mealy machine is invertible if all functions $\varphi_k$ are bijections.

A cryptomachine $C = (X, Q, Y, K, \psi, \varphi)$ is called a Zakrevs'kij cipher (see [2]) if the set $\mathcal{Y} = \{C(q_0, k) | (q_0, k) \in Q \times K_0\}$ is the set of all invertible Mealy machines with strongly connected underlying digraphs. In particular, this means that $|X| = |Y|$.

Let $\mathcal{A} = (X, Q, Y, K, \psi, \varphi)$ be a Mealy machine. The transition function $\psi$ can be extended in a unique way to an action $Q \times X^* \rightarrow Q$. We still denote this extension by $\psi$. The output function $\varphi$ can also be inductively extended to an action $Q \times X^* \rightarrow Y^*$ if for all $w \in X^*$, $a \in X$, and $q \in Q$ we set $\varphi(q, w, a) = \varphi(q, w, \psi(q, w, a))$.

Thus every Mealy machine generates the word transformation function $\varphi_w : X^* \rightarrow Y^*$ (where $\varphi_w(\omega) = \varphi(q_0, \omega)$ for every $\omega \in X^*$). We say that two Mealy machines are equivalent if their word transformation functions coincide. If we take two equivalent Mealy machines from the set $\mathcal{Y}$, then we say that the corresponding keys are equivalent.

There is a natural question: how many nonequivalent keys are there in a Zakrevs'kij cipher with given number of states and given number of letters? Another formulation of this question is as follows: how many nonequivalent invertible Mealy automata with strongly connected underlying digraphs are there?

As shown in [1], if $|Q| = n$ and $|X| = |Y| = m$, then the number of nonequivalent Zakrevs'kij cipher keys is at most $(mn)^{2m}/(n-1)!$ and at least $m^n$. We present a new lower bound for this number.

2. The lower bound

2.1. Notation. For a word $w \in \{a, b\}^*$, we denote by $|w|$ the length of $w$ and by $w[i]$, where $1 \leq i \leq |w|$, the $i$th letter in $w$ from the left. If $1 \leq i \leq j \leq |w|$, we denote by $w[i, j]$ the word $w[i] \cdots w[j]$.

Throughout the paper, we assume that the state set $Q$ of the automata under consideration is the set $\{0, 1, 2, \ldots, n-1\}$ of the first $n$ nonnegative integers.

We denote by $a \mod b$ the integer $x$ such that $0 \leq x < b$, $b \mid (a - x)$.

2.2. The lower bound

Theorem 2.1. Let $n, m \geq 2$, $n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, where $p_i$ are pairwise distinct primes, $\alpha_i \geq 1$. Then there are at least

$$n^{n(m-1)} \cdot (m!^{p_1-1}) \cdots (m!^{p_k-1}) \cdot m^{n-p_1-p_2-\cdots-p_k+k}$$

pairwise nonequivalent keys in a Zakrevs'kij cipher over automata with $n$ states and $m$-letter alphabet.
Proof. Consider the set $\mathcal{Y}_1$ of all Mealy machines with $n$ states, input alphabet $X$, and output alphabet $Y$ of the following structure:

- $q_0 = 0$;
- for every $i < n - 1$, $\varphi(i, a) = i + 1$; $\varphi(n - 1, a) = 0$;
- for every $i \in \{1, 2, \ldots, k\}$ there exist $j \in \{1, \ldots, p_i - 1\}, c \in X$ such that $\varphi(\frac{j c}{p_i}, c) \neq \varphi(0, c)$;
- for every $i \in Q$, for any $c, d \in X, c \neq d \Rightarrow \varphi(i, c) \neq \varphi(i, d)$.

That is, the letter $a$ from the input alphabet $X$ is special and sends the automaton to the next state in the linear order. For all other transitions, the restrictions are weaker and we only require that if we choose any prime divisor $p$ of $n$ and consider the states $0, \frac{p}{p}, \frac{2p}{p}, \ldots, \frac{(p - 1)p}{p}$, then not all of them have the same output function.

Note that the transitions by the letter $a$ make the underlying graphs of all these automata strongly connected, and the restrictions on the output function $\varphi$ ensure that every automaton from $\mathcal{Y}_1$ is invertible, that is, $\mathcal{Y}_1 \subseteq \mathcal{Y}$.

Choose arbitrary automata $\mathcal{A}_1 \neq \mathcal{A}_2 \in \mathcal{Y}_1$. We will prove that they are nonequivalent.

Let $\mathcal{A}_1 = \langle X, Q, Y, \psi_1, \varphi_1 \rangle$, $\mathcal{A}_2 = \langle X, Q, Y, \psi_2, \varphi_2 \rangle$. There are two possible cases:

**Case 1.** There exist $i \in Q, c \in X$ such that $\varphi_1(i, c) \neq \varphi_2(i, c)$, that is, the output functions differ in some state $i$. Consider the words $\omega_1 d_1 = \varphi_1(0, a^c)$ and $\omega_2 d_2 = \varphi_2(0, a^c)$. Note that they differ in the last letter, because $d_1 = \varphi_1(i, c) \neq \varphi_2(i, c) = d_2$.

**Case 2.** For any $q \in Q, b \in X$, we have $\varphi_1(q, b) = \varphi_2(q, b) = \varphi(q, b)$. Then $\psi_1 \neq \psi_2$, that is, there is a state $e \in Q$ and a letter $c \in X$ such that $\psi_1(e, c) \neq \psi_2(e, c)$. Let $\psi_1(e, c) = q_1, \psi_2(e, c) = q_2$.

Set $d (p_1, p_2) = \min \{i \mid \psi(q_1, a^t) = p_2\}$. Note that this notation is correct for any $p_1, p_2 \in Q$, because the transitions by the letter $a$ produce a cycle over all states of the automaton. In what follows, we call this function the distance between a pair of states. It is easy to see that in our notation, $d(p_1, p_2) = (p_2 - p_1) \mod n$ and $d(\psi(q_1, a), \psi(q_2, a)) = d(p_1, p_2)$, that is, transitions by the letter $a$ do not change the distances between states.

Let $d(q_1, q_2) = t$. If we consider

$$\psi(q_1, a^0), \psi(q_2, a^0), (\psi(q_1, a^t), \psi(q_2, a^t)), \ldots, (\psi(q_1, a^{t^n}), \psi(q_2, a^{t^n})), \ldots$$

then we obtain all pairs of states with distance $t$ between them. If we recall that $q_1 = \psi_1(0, a^c), q_2 = \psi_2(0, a^c)$, this means that for any pair of states $p_1, p_2$, if $d(p_1, p_2) = t$ then there is a word $\omega_1 = a^tca^t$ such that $\psi_1(0, \omega_1) = p_1, \psi_2(0, \omega_1) = p_2$.

Consider the set $Q_1 = \{0, t \mod n, 2t \mod n, \ldots, n \cdot t \mod n = 0\}$.

**Proposition 2.2.** There exist $p_1, p_2 \in Q_1, b \in X$ such that $\varphi(p_1, b) \neq \varphi(p_2, b)$.

**Proof.** Let $\gcd(n, t) = d$. According to Bezout’s identity,

$$x y \in \mathbb{Z} \text{ such that } 0 \leq x < n, \quad t \cdot x + n \cdot y = d.$$

Choose any prime $p$ such that $p \mid \frac{n}{d}$. Then $d \mid \frac{n}{p}$, so that

$$x_1, y_1 \in \mathbb{Z} \text{ such that } 0 \leq x_1 < n, \quad t \cdot x_1 + n y_1 = \frac{n}{p};$$

therefore, $x_1 \cdot t \mod n = \frac{n}{p}$, whence $\frac{n}{p} \in Q_1$. Analogously, for every $i \in \{0, 1, \ldots, p - 1\}$ we have $\frac{mi}{p} \in Q_1$. By the definition of $\varphi$,

there exist $j \in \{0, \ldots, p - 1\}, b \in X$ such that $\varphi(\frac{nj}{p}, b) \neq \varphi(0, b)$. \hfill \Box

Consider $p_1, p_2 \in Q_1$ from Proposition 2.2. Let

$$p_1 = i \cdot t \mod n, \quad p_2 = j \cdot t \mod n.$$

Without loss of generality, $i < j$. Since

$$\varphi(i \cdot t \mod n, b) \neq \varphi(j \cdot t \mod n, b),$$

there is $l \in \mathbb{Z}$ such that $i \leq l < j, \varphi(l \cdot t \mod n, b) \neq \varphi((l + 1) \cdot t \mod n, b)$.

We have $d(l \cdot t \mod n, (l + 1) \cdot t \mod n) = t$; therefore, there exists $s \in \{0, 1, \ldots, n - 1\}$ such that $\psi_1(0, a^sca^t) = l \cdot t \mod n$ and $\psi_2(0, a^sca^t) = (l + 1) \cdot t \mod n$, whence $\varphi_1(0, a^sca^t b) \neq \varphi_2(0, a^sca^t b)$.

Thus any two automata from the set $Y_1$ are nonequivalent; therefore, the total number of nonequivalent keys of a Zakrevskij cipher is at least $|Y_1|.$
The cardinality of $Y_1$ is the number of admissible pairs of a transition function and an output function. The number of admissible transition functions is equal to $n^{n(m-1)}$, because for any state $q \in Q$ and for any letter $b \in X$ except $a$, we can choose an arbitrary $p \in Q$ as the value $\psi(q, b)$.

In any state $q \in Q$, we can choose any of the $m!$ bijections $X \to Y$ as the output function in this state. But we have a restriction: for any set $\{\frac{n}{m}, \frac{2n}{m}, \ldots, \frac{(p-1)n}{m}\}$, the output function in at least one of its elements must differ from the output function in the first state. Thus there are

$$(m!^{|p|_1 - 1} - 1) \cdot (m!^{|p_2| - 1} - 1) \cdots (m!^{|p_k| - 1} - 1)m!^{n-p_1-p_2-\cdots-p_k}$$

admissible output functions.

So,

$$|Y_1| = n^{n(m-1)}(m!^{p_1 - 1} - 1) \cdot (m!^{p_2 - 1} - 1) \cdots (m!^{p_k - 1} - 1) \cdot m!^{n-p_1-p_2-\cdots-p_k + k}.$$  

\[\square\]

REFERENCES