МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

UDC 519.7 DOI 10.17223/20710410/46/3

CRYPTANALYTICAL FINITE AUTOMATON INVERTIBILITY WITH FINITE DELAY[1]

G. P. Agibalov

National Research Tomsk State University, Tomsk, Russia

E-mail: agibalov@mail.tsu.ru

The paper continues an investigation of the cryptanalytical invertibility concept with a finite delay introduced by the author for finite automata. Here, we expound an algorithmic test for an automaton A to be cryptanalytically invertible with a finite delay, that is, to have a recovering function f which allows to calculate a prefix of a length m in an input sequence of the automaton A by using its output sequence of a length $m + \tau$ and some additional information about A defining a type of its invertibility and known to cryptanalysts. The test finds out whether the automaton A has a recovering function f or not and if it has, determines some or, may be, all of such functions. The test algorithm simulates a backtracking method for searching a possibility to transform a binary relation to a function by shortening its domain to a set corresponding to the invertibility type under consideration.

Keywords: *finite automata, information-lossless automata, automata invertibility, cryptanalytical invertibility, cryptanalytical invertibility test.*

Introduction

To continue the research we have begun in [1], we first present the problem under consideration, namely the automaton cryptanalytical invertibility, and connected with it basic concepts and terms.

An arbitrary finite automaton is represented by a 5-tuple $A = (X, Q, Y, \psi, \varphi)$, where X, Q, and Y are the input alphabet, the set of states and the output alphabet respectively, $\psi : X \times Q \to Q$ and $\varphi : X \times Q \to Y$. The last functions, being defined for pairs $xq \in X \times Q$, are expanded on pairs $\alpha q \in X^* \times Q$ by induction on the length $|\alpha|$ of a word $\alpha \in X^*$, namely the functions $\psi : X^* \times Q \to Q$ and $\bar\varphi : X^* \times Q \to Y^*$ are defined as $\psi(\Lambda, q) = q$, $\psi(\alpha\beta, q) = \psi(\beta, \psi(\alpha, q))$, $\bar\varphi(\Lambda, q) = \Lambda$, $\bar\varphi(x, q) = \varphi(x, q)$ and $\bar\varphi(\alpha\beta, q) = \bar\varphi(\alpha, q)\bar\varphi(\beta, \psi(\alpha, q))$. The symbol Λ here denotes the empty word in any alphabet. Thus, $\psi(\alpha, q)$ is a state to which the automaton A goes from the state q under the action of the input word α, and $\bar\varphi(\alpha, q)$ is a word which it outputs under this action.

Everywhere further, τ means a natural number and is called a finite delay, and without another note, it is supposed that $\alpha \in X^m$ for $m = |\alpha\delta| - \tau$, $\delta \in X^\tau$, $q \in Q$. In dependence on context, the last symbols are considered as elements of the pointed sets respectively or as variables with these sets as their ranges.

In connection with the automaton A, we believe that q, α, and δ are the variables with values from Q, X^m, and X^τ denoting, respectively, an initial state, an information word, and

[1]The author is supported by the RFBR-grant no. 17-01-00354.

a delay word in an input sequence $\alpha\delta$ of the automaton A, and $K = \{\forall q, \forall\alpha, \forall\delta, \exists q, \exists\delta\}$ is a set of universal and existential quantifiers that binds these variables. Note that in K there is no the quantifier $\exists\alpha$. This is explained with the following argument: for a cryptanalyst, an information word α in an input word of the automaton A is supposed to be unknown and not some one, but any one. As for the length $m = |\alpha|$ of the word α, it is proposed to be known since it can be calculated as it is shown above where $|\alpha\delta| = |\bar{\varphi}(\alpha\delta, q)|$ and $\bar{\varphi}(\alpha\delta, q)$ is a sequence supervised on the output of A by a cryptanalyst. Also, let $V_0 = \{q, \delta, \psi(\alpha, q), \psi(\alpha\delta, q)\}$ where q, $\psi(\alpha, q)$, and $\psi(\alpha\delta, q)$ are, respectively, the initial, intermediate and final states of the automaton A and δ is a delay word. For any subset $v \subseteq V_0$, let $v(q, \alpha, \delta)$ be the system of functions (or vector function) represented by the formulas in v and depending on variables q, α, δ denoting respectively an initial state, an input word, and a delay word in the automaton A. Denote D_v the range of the function $v(q, \alpha, \delta)$, that is, the set of its possible values.

1. Automata cryptanalitical invertibility problem

The automaton A is called (cryptanalitically) invertible with a delay τ if there exist quantifiers K_1, K_2, K_3 in K with the different variables from $\{q, \alpha, \delta\}$, a subset $v \subseteq V_0$ and a function $f : Y^{m+\tau} \times D_v \to X^m$ such that

$$K_1K_2K_3(f(\bar{\varphi}(\alpha\delta, q), v(q, \alpha, \delta)) = \alpha); \tag{1}$$

in this case f is called a recovering function (it recovers α using $\bar{\varphi}(\alpha\delta, q)$ and $v(q, \alpha, \delta)$), the 4-tuple $IT = (K_1K_2K_3, v)$, the triple $ID = (K_1K_2K_3)$, and $IO = v$ are respectively called a type, a degree, and an order of (cryptanalytical) invertibility of the automaton A.

In this definition, $K_i = Q_ix_i$ for each $i = 1, 2, 3$, a quantifier symbol $Q_i \in \{\forall, \exists\}$, and a variable $x_i \in \{q, \alpha, \delta\}$. Therefore, at the same time in the future, we equally use (1) and the expression

$$Q_1x_1Q_2x_2Q_3x_3(f(\bar{\varphi}(\alpha\delta, q), v(q, \alpha, \delta)) = \alpha), \tag{2}$$

where $\{x_1, x_2, x_3\} = \{q, \alpha, \delta\}$ and $Q_1x_1Q_2x_2Q_3x_3 = ID$.

The main problem that we consider in the paper, the problem of automata cryptanalytical invertibility — ACI, is the following decision one: given a finite automaton $A = (X, Q, Y, \psi, \varphi)$, an invertibility type $IT = (K_1K_2K_3, v) = (Q_1x_1Q_2x_2Q_3x_3, v)$, and a natural number τ, find out whether the automaton A is invertible of type IT with the delay τ and if so, construct a proper recovering function f satisfying the any of conditions (1) or (2).

2. Function cryptanalytical invertibility problem

To decide the problem ACI, we first try to decide the following auxiliary abstract mathematical problem of function invertibility — FI: given a function $g(x_1, \ldots, x_n)$, a quantifier word $Q_1x_1 \ldots Q_nx_n$, and a number $k_0 \in \{1, \ldots, n\}$ where $Q_{k_0} = \forall$, find out if there exist functions f such that the formula

$$Q_1x_1Q_2x_2 \ldots Q_nx_n(f(g(x_1, x_2, \ldots, x_n)) = x_{k_0}) \tag{3}$$

is true, and if exist, construct some of them.

Using the terms related to the cryptanalytical invertibility of an automaton, we can say that in this problem the question is about the invertibility of type $(Q_1x_1 \ldots Q_nx_n)$ for the function $g(x_1, \ldots, x_n)$ with respect to a variable x_{k_0} and with a recovering function $f : D_g \to D_{k_0}$ where D_g and D_{k_0} are the ranges of the function g and of the variable x_{k_0} respectively.

Clearly, the main problem (ACI) is obtained from the auxiliary one (FI) as the following particular case: $n = 3$, $k_0 \in \{1,2,3\}$, $\{x_1, x_2, x_3\} = \{q, \alpha, \delta\}$, $x_{k_0} = \alpha$, $g(x_1, x_2, x_3) = (\bar{\varphi}(\alpha\delta, q), \upsilon(q, \alpha, \delta))$. Thus, any method deciding the auxiliary problem also decides the main one, and, so, our problem ACI reduces to our problem FI.

Every of the predicate logic formulas under consideration in the paper including (1)–(3) is written in a normal form $Q_1x_1 \ldots Q_nx_nP(x_1, \ldots, x_n)$, that is, with a quantifier prefix $Q_1x_1 \ldots Q_nx_n$ and its scope being an underlying predicate expression $P(x_1, \ldots, x_n)$ without quantifiers and, moreover, of the special kind $(f(g(x_1, \ldots, x_n)) = x_{k_0})$. We consider the quantifier prefix $Q_1x_1 \ldots Q_nx_n$ in it as a way to define a domain of values of subject variables $x_1, \ldots, x_n$ that the underlying function $P(x_1, \ldots, x_n)$ depends on. In fact, the quantifier prefix in it generates some n-tuples $a = a_1 \ldots a_n$ of values for the variable $x = x_1 \ldots x_n$, and the underlying expression calculates the values $g(a)$ and determines f by the equalities $f(g(a)) = a_{k_0}$. According to the quantifier logic [2], the quantifier $\forall x_k$ generates all the possible values a_k of the variable x_k from its range D_k and the quantifier $\exists x_k$ generates a one of the possible values a_k taken from D_k in dependence on the values $a_1, \ldots, a_{k-1}$ of the previous variables $x_1, \ldots, x_{k-1}$ respectively. From the cryptanalytical point of view, we suppose that the value a_k provided by the quantifier $\exists x_k$ as well as the rule of its generating, the function $h(x_1, \ldots, x_n) = f(g(x_1, \ldots, x_n))$ and, in general, the function of $P(x_1, \ldots, x_n)$ are a priori unknown to a cryptanalyst.

Note that under suppositions named above, we are forced to decide the FI problem by trying different values allegedly generated by an existential quantifier what many times complicates the deciding algorithm. The same effect results from determining a function f by the equations $f(g(a)) = a_{k_0}$, because the last very often (for example, when $g(a) = g(b)$ and $a_{k_0} \neq b_{k_0}$) determines not function f but a binary relation f which is not a function.

Consider (3) taking into account that has been just said in relation to the FI problem. Let $n = r + s$, $r \geqslant 1$, $s \geqslant 0$, $i_1 < \ldots < i_r$, $j_1 < \ldots < j_s$, $\{i_1 \ldots, i_r, j_1, \ldots, j_s\} = \{1, \ldots, n\}$, $Q_{i_1} = \ldots = Q_{i_r} = \forall$, $Q_{j_1} = \ldots = Q_{j_s} = \exists$, $D_1, \ldots, D_n$ and D_g are the ranges of variables x_1, $\ldots$, x_n and the function g respectively. So $k_0 \in \{i_1, \ldots, i_r\}$, $Q_{k_0} = \forall$, $g : D_1 \times \ldots \times D_n \to D_g$, $f : D_g \to D_{k_0}$. In the case $s = 0$ it is supposed that $\{j_1, \ldots, j_s\} = \varnothing$. Also, let for $k \in \{j_1, \ldots, j_s\}$, $\varepsilon_k : D_1 \times \ldots \times D_{k-1} \to D_k$ and $\varepsilon_k(a_1, \ldots, a_{k-1})$ denotes a value of the variable x_k, the existence of which is implied by a quantifier Q_kx_k with $Q_k = \exists$ in dependence on the values $a_1, \ldots, a_{k-1}$ chosen before by the quantifiers $Q_1x_1, \ldots, Q_{k-1}x_{k-1}$ for the variables $x_1, \ldots, x_{k-1}$ respectively. Further, in order to address or refer to functions $\varepsilon_k(a_1, \ldots, a_{k-1})$, we call them existential ones for their relation to quantifiers of the similar name. A function $\varepsilon_k(a_1, \ldots, a_{k-1})$ isn't obliged to essentially depend on each of its arguments. In this case we exclude inessential arguments from the list under the sign of the function. At last, if $s = 0$, that is, in the quantifier prefix under consideration there are no existential quantifiers and hence $D_{j_1} \times \ldots \times D_{j_s} = \varnothing$, then we have $\varepsilon_1 \ldots \varepsilon_s = \Lambda$.

Believing the value $\varepsilon_k(a_1, \ldots, a_{k-1})$ be unknown, to find out it we can try different elements a_k in D_k as the real value for ε_k and to pick out that of them, for which the equations $f(g(a_1, \ldots, a_n)) = a_{k_0}$ determine f as a function. In the case when no element in D_k satisfies this condition, we can change the value a_{k-1} of the previous variable x_{k-1} like in the method of backtracking search tree traversal [3–5].

For the quantifier prefix $Q_1x_1 \ldots Q_nx_n$ in (3), define a subset $M_n \subseteq D_1 \times \ldots \times D_n$ by induction on $k = 1, 2, \ldots, n$, namely let $M_0 = \{\Lambda\}$ and for each $k \in \{1, \ldots, n\}$, if $k \in \{i_1, \ldots, i_r\}$, then $M_k = \{a_1 \ldots a_{k-1}a_k : a_1 \ldots a_{k-1} \in M_{k-1}, a_k \in D_k\} = M_{k-1} \times D_k$, otherwise if $k \in \{j_1, \ldots, j_s\}$, then $M_k = \{a_1 \ldots a_{k-1}a_k : a_1 \ldots a_{k-1} \in M_{k-1}, a_k = \varepsilon_k(a_1, \ldots, a_{k-1})\} = M_{k-1} \times \{\varepsilon_k(a_1, \ldots, a_{k-1})\}$. By this definition, M_n is uniquely defined by the existential functions $\varepsilon_k(a_1,\ldots,a_{k-1})$, $k\in\{j_1,\ldots,j_s\}$. Therefore, we denote it M_ε where $\varepsilon=\varepsilon_{j_1}\ldots\varepsilon_{j_s}$ is the vector existential function of $Q_1x_1\ldots Q_nx_n$, say that M_ε corresponds to these functions or shorter to ε and call M_ε the existential domain of the predicate word $Q_1x_1\ldots Q_nx_n$ corresponding to existential functions in ε.

Notice that by the definition,

$$a_1\ldots a_n\in M_\varepsilon\Leftrightarrow$$
$$\Leftrightarrow(a_1\ldots a_n\in D_1\times\ldots\times D_n)\;\&\;(a_{j_1}\ldots a_{j_s}=\varepsilon_{j_1}(a_1,\ldots,a_{j_1-1})\ldots\varepsilon_{j_s}(a_1,\ldots,a_{j_s-1})),$$

that is, M_ε consists of those vectors $a_1\ldots a_n$ in $D_1\times\ldots\times D_n$ which are generated by the quantifier prefix by means of existential functions $\varepsilon_{j_1},\ldots,\varepsilon_{j_s}$ (independently of the underlying expression) in such a way that a_k is any element in D_k if $Q_k=\forall$ or it is $\varepsilon_k(a_1,\ldots,a_{k-1})$ otherwise, $k\in\{1,\ldots,n\}$.

Also, please pay attention to the following property of the set M_ε, resulting from the functionality of mappings ε_k in its definition: for all $a_1a_2\ldots a_n$ and $b_1b_2\ldots b_n$ in M_ε and for any $k\in\{j_1,\ldots,j_s\}$ if $a_1\ldots a_{k-1}=b_1\ldots b_{k-1}$, then $a_k=\varepsilon_k(a_1,\ldots,a_{k-1})=$ $=\varepsilon_k(b_1,\ldots,b_{k-1})=b_k$.

Further, in dependence on context, we use the terms of existential function $\varepsilon_k(a_1,\ldots,a_{k-1})$ and of existential domain M_ε in connection not only with a quantifier prefix $Q_1x_1\ldots Q_nx_n$ but with an automaton invertibility degree being denoted in the same way.

Now, we give some examples demonstrating what we have just discussed.

3. Examples of existential functions and domains of a predicate prefix

Example 1. Let $n=3$, $x=x_1x_2x_3$, $g(x)=g(x_1,x_2,x_3)=(x_1x_2+x_3)\bmod 3$, $D_1=D_2=D_3=D_g=\{0,1,2\}$, $k_0=1$ and $x_{k_0}=x_1$, $f:D_g\to D_1$, $Q_1x_1Q_2x_2Q_3x_3(f(g(x_1,x_2,x_3))=x_{k_0})=\forall x_1\forall x_2\exists x_3(f((x_1x_2+x_3)\bmod 3)=x_1)$, the function $\varepsilon_3:D_1\times D_2\to D_3$ is given in the Table 1.

Table 1

x_1x_2	00	01	02	10	11	12	20	21	22
$\varepsilon_3(x_1,x_2)$	2	2	2	1	0	2	0	1	2

Then $M_\varepsilon=M_{\varepsilon_3}=\{002,012,022,101,110,122,200,211,222\}$, the values $g(x)$ and $f(g(x))$ for $x\in M_\varepsilon$ are presented in the Table 2.

Table 2

$x\in M_\varepsilon$	002	012	022	101	110	122	200	211	222
$g(x)$	2	2	2	1	1	1	0	0	0
$f(g(x))$	0	0	0	1	1	1	2	2	2

It is immediately seen that M_ε is really generated by the quantifier prefix $\forall x_1\forall x_2\exists x_3$ by means of the existential function ε_3, and f is a function on D_g with the values in D_1, satisfying the underlying predicate equation for vectors in M_ε and hence proving the invertibility of type $(\forall x_1,\forall x_2,\exists x_3)$ of the function g with respect to the variable x_1 and with the recovering function f. We can add that in fact there are yet at least five other existential functions and five other recovering functions f, with which the function g in the example is invertible of the type $\forall x_1\forall x_2\exists x_3$ with respect to the variable x_1.

Example 2. This example only differs from the first one in the range D_3 which now is $D_3 = \{0, 1\}$ and in the existential function $\varepsilon_3 : D_1 \times D_2 \to D_3$ (Table 3).

Table 3

x_1x_2	00	01	02	10	11	12	20	21	22
$\varepsilon_3(x_1, x_2)$	0	0	0	1	0	1	1	0	1

In this case we obtain the following set

$$M_\varepsilon = M_{\varepsilon_3} = \{000, 010, 020, 101, 110, 121, 201, 210, 221\},$$

and the following functions $g(x)$ and $f(g(x))$ defined on it (Table 4).

Table 4

$x \in M_\varepsilon$	000	010	020	101	110	121	201	210	221
$g(x)$	0	0	0	1	1	0	1	2	2
$f(g(x))$	0	0	0	1	1	0	1	2	2

From the Table 4, it is seen that f is a function on D_g but it doesn't satisfy the equation $f(g(x_1, x_2, x_3)) = x_1$ on M_ε and therefore g is not invertible of the type $\forall x_1 \forall x_2 \exists x_3$ with existential function ε and with respect to the variable x_1. There is a suspicion that it is not invertible of this type with any existential function ε for $\exists$ and with respect to the same variable.

4. Existential functions and domains of automaton invertibility degrees

In [1], all the possible automaton cryptanalytical invertibility types were defined. In the section 1 of this paper, we have repeated the definition. Each type IT is characterised by an invertibility degree ID and invertibility order IO. Here, for each of all thirteen possible IDs $Q_1x_1Q_2x_2Q_3x_3$ of an automaton $A = (X, Q, Y, \psi, \varphi)$, we give the general description of ranges and domains for arbitrary existential functions $\varepsilon_1, \varepsilon_2, \varepsilon_3$ in it and, for any $\varepsilon \subseteq \{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$, the general description of existential domain M_ε in the form of algorithm for computing vectors from $D_1 \times D_2 \times D_3$ in it.

In order to make the text of this section to be nearer to the automata theory language which we keep to in our research, instead of typical symbols x_1, x_2, and x_3 of abstract mathematical variables, we use the symbols q, α, and δ usually denoting in automata theory an initial state of the automaton A, its input information and its input delay words respectively. Besides, m is the length of α and τ is the length of δ.

1) $ID = \forall q \forall \alpha \forall \delta$, $\varepsilon = \Lambda$, $M_\varepsilon = \{q\alpha\delta : q \in Q, \alpha \in X^m, \delta \in X^\tau\}$;
2) $ID = \forall q \forall \alpha \exists \delta$, $\varepsilon_3 : Q \times X^m \to X^\tau$, $M_\varepsilon = M_{\varepsilon_3} = \{q\alpha\varepsilon_3(q, \alpha) : q \in Q, \alpha \in X^m\}$;
3) $ID = \forall q \exists \delta \forall \alpha$, $\varepsilon_2 : Q \to X^\tau$, $M_\varepsilon = M_{\varepsilon_2} = \{q\varepsilon_2(q)\alpha : q \in Q, \alpha \in X^m\}$;
4) $ID = \exists q \forall \alpha \forall \delta$, $\varepsilon_1 \in Q$, $M_\varepsilon = M_{\varepsilon_1} = \{\varepsilon_1\alpha\delta : \alpha \in X^m, \delta \in X^\tau\}$;
5) $ID = \exists q \forall \alpha \exists \delta$, $\varepsilon_1 \in Q$, $\varepsilon_3 : Q \times X^m \to X^\tau$, $M_\varepsilon = M_{\varepsilon_1\varepsilon_3} = \{\varepsilon_1\alpha\varepsilon_3(\varepsilon_1, \alpha) : \alpha \in X^m\}$;
6) $ID = \exists q \exists \delta \forall \alpha$, $\varepsilon_1 \in Q$, $\varepsilon_2 : Q \to X^\tau$, $M_\varepsilon = M_{\varepsilon_1\varepsilon_2} = \{\varepsilon_1\varepsilon_2(\varepsilon_1)\alpha : \alpha \in X^m\}$;
7) $ID = \forall \alpha \exists q \forall \delta$, $\varepsilon_2 : X^m \to Q$, $M_\varepsilon = M_{\varepsilon_2} = \{\alpha\varepsilon_2(\alpha)\delta : \alpha \in X^m, \delta \in X^\tau\}$;
8) $ID = \forall \alpha \exists q \exists \delta$, $\varepsilon_2 : X^m \to Q$, $\varepsilon_3 : X^m \times Q \to X^\tau$, $M_\varepsilon = M_{\varepsilon_2\varepsilon_3} =$
$= \{\alpha\varepsilon_2(\alpha)\varepsilon_3(\alpha, \varepsilon_2(\alpha)) : \alpha \in X^m\}$;
9) $ID = \forall \alpha \forall \delta \exists q$, $\varepsilon_3 : X^m \times X^\tau \to Q$, $M_\varepsilon = M_{\varepsilon_3} = \{\alpha\delta\varepsilon_3(\alpha, \delta) : \alpha \in X^m, \delta \in X^\tau\}$;
10) $ID = \forall \alpha \exists \delta \forall q$, $\varepsilon_2 : X^m \to X^\tau$, $M_\varepsilon = M_{\varepsilon_2} = \{\alpha\varepsilon_2(\alpha)q : \alpha \in X^m, q \in Q\}$;

11) $ID = \forall\delta\exists q\forall\alpha$, $\varepsilon_2 : X^\tau \to Q$, $M_\varepsilon = M_{\varepsilon_2} = \{\delta\varepsilon_2(\delta)\alpha : \alpha \in X^m, \delta \in X^\tau\}$;
12) $ID = \exists\delta\forall q\forall\alpha$, $\varepsilon_1 \in X^\tau$, $M_\varepsilon = M_{\varepsilon_1} = \{\varepsilon_1 q\alpha : q \in Q, \alpha \in X^m\}$;
13) $ID = \exists\delta\forall\alpha\exists q$, $\varepsilon_1 \in X^\tau$, $\varepsilon_3 : X^\tau \times X^m \to Q$, $M_\varepsilon = M_{\varepsilon_1\varepsilon_3} = \{\varepsilon_1\alpha\varepsilon_3(\varepsilon_1, \alpha) : \alpha \in X^m\}$.

From the given expressions for the sets M_ε, we can see the expressions for the size $|M_\varepsilon|$ of these sets. The Table 5 contains them for all numbers of ID. In it $k = |X|$, $h = |Q|$.

Table 5

$N.ID$	1	2	3	4	5	6	7	8	9	10	11	12	13
$\lvert M_\varepsilon\rvert$	$hk^{m+\tau}$	hk^m	hk^m	$k^{m+\tau}$	k^m	k^m	$k^{m+\tau}$	k^m	$k^{m+\tau}$	hk^m	$k^{m+\tau}$	hk^m	k^m

5. Test for function cryptanalytical invertibility

Lemma 1. For a function $g(x_1, \ldots, x_n)$, there exists a function $f : D_g \to D_{k_0}$ with the true formula (3), if and only if for each $k \in \{j_1, \ldots, j_s\}$ there exists an existential function $\varepsilon_k : D_1 \times \ldots \times D_{k-1} \to D_k$ such that the set M_ε corresponding to $\varepsilon = \varepsilon_{j_1} \ldots \varepsilon_{j_s}$ satisfies the following condition:

$$\forall a = a_1 \ldots a_n \in M_\varepsilon \, \forall b = b_1 \ldots b_n \in M_\varepsilon \, (a_{k_0} \neq b_{k_0} \Rightarrow g(a) \neq g(b)). \tag{4}$$

Proof. Necessity: given $\exists f((3))$, prove $\exists\varepsilon((4))$. We have:

$$\begin{gathered}\exists f((3)) \Rightarrow \exists f(Q_1x_1 \ldots Q_nx_n(f(g(x_1, \ldots, x_n)) = x_{k_0})) \Rightarrow \\ \Rightarrow \exists f\exists\varepsilon(\forall a = a_1 \ldots a_n \in M_\varepsilon(f(g(a)) = a_{k_0})) \Rightarrow \\ \Rightarrow \exists f\exists\varepsilon(\forall a \in M_\varepsilon(f(g(a)) = a_{k_0}) \ \& \ \forall b \in M_\varepsilon(f(g(b)) = b_{k_0})) \Rightarrow \\ \Rightarrow \exists f\exists\varepsilon(\forall a \in M_\varepsilon \, \forall b \in M_\varepsilon(f(g(a)) = a_{k_0}) \ \& \ (f(g(b)) = b_{k_0})) \Rightarrow \\ \Rightarrow \exists f\exists\varepsilon(\forall a \in M_\varepsilon \, \forall b \in M_\varepsilon(a_{k_0} \neq b_{k_0} \Rightarrow f(g(a)) \neq f(g(b)))) \Rightarrow \\ \Rightarrow \exists\varepsilon(\forall a \in M_\varepsilon \, \forall b \in M_\varepsilon(a_{k_0} \neq b_{k_0} \Rightarrow g(a) \neq g(b))) = \exists\varepsilon((4)).\end{gathered}$$

Sufficiency: given $\exists\varepsilon((4))$, prove $\exists f((3))$. Define $f : D_g \to D_{k_0}$ as $f(g(a)) = a_{k_0}$, $a \in M_\varepsilon$. We have:

$$\exists\varepsilon((4)) = \exists\varepsilon(\forall a \in M_\varepsilon \, \forall b \in M_\varepsilon(a_{k_0} \neq b_{k_0} \Rightarrow g(a) \neq g(b)).$$

Therefore, $g(a) = g(b) \Rightarrow a_{k_0} = b_{k_0} \Rightarrow f(g(a)) = f(g(b))$ for any a and b from M_ε what means that f is a function on $\{g(a) : a \in M_\varepsilon\}$. So, $\forall a \in M_\varepsilon(f(g(a)) = a_{k_0})$ that is equivalent to $((3))$. ■

So, by trying the different existential functions ε on satisfying the existential domains M_ε to the condition (4), we can find out whether there exists a function f recovering a certain variable of a given function g or not.

Corollary 1. A function $g(x_1, \ldots, x_n)$ is invertible of a type $Q_1x_1 \ldots Q_nx_n$ with respect to a variable x_{k_0}, $k_0 \in \{i_1, \ldots, i_r\}$, if and only if there exist existential functions $\varepsilon_k : D_1 \times \ldots \times D_{k-1} \to D_k$, $k = j_1, \ldots, j_s$, the corresponding to which set M_ε satisfies the condition (4).

So, by trying the different existential functions ε on satisfying the existential domains M_ε to the condition (4), we can find out whether a function g is invertible of a certain type with respect to some its variable or not.

Example 3. This is the end of Example 1. We see here that $\forall x_1\forall x_2\exists x_3(f(g(x)) = x_1)$, that is, the state (3) is true as well as $\forall a \in M_\varepsilon \, \forall b \in M_\varepsilon(a_1 \neq b_1 \Rightarrow g(a) \neq g(b))$, that is, the state (4) is true too. For instance, if $(x_1, x_2) = (1, 2)$, then $x_3 = 2$, $g(x_1, x_2, x_3) = (x_1x_2 + x_3) \bmod 3 = 1$ and $f(g(x_1, x_2, x_3)) = f(1) = 1 = x_1$, and also if $a = 020$ and $b = 101$, then $a_1 = 0 \neq 1 = b_1$ and $g(a) = g(020) = 0 \neq 1 = g(101) = g(b)$.

Example 4. This is the end of example 2. Here, both conditions (3) and (4) are false because, for example, $f(g(x)) \neq x_1$ for $x = 121$ and $x = 201$, $a_1 \neq b_1$ and $g(a) = g(b)$ for $a = 000$ and $b = 121$, for $a = 020$ and $b = 121$. Moreover, immediately from the Table 4, it is seen that, for this g, there doesn't exist f with the property $f(g(x)) = x_1$ and it isn't possible to recover the value x_1 from the value $g(x)$. Also, it's impossible to make the condition (4) to become true in the way of choosing other values for x_3 in points $x \in M_\varepsilon$, since for any values a_3, b_3 of variable x_3, there exist some values a_1, a_2 and b_1, b_2 of the variables x_1, x_2 such that a_1, a_2 are arbitrary, b_1 is invertible modulo 3 and $b_2 = b_1^{-1}(a_1a_2 + a_3 - b_3) \bmod 3$ and then we will have what we need, namely: $a_1 \neq b_1$ and $g(a) = a_1a_2 + a_3 = b_1b_2 + b_3 = g(b)$. So, if for x we take the value $b' = 120$ instead of $b = 121$, then, from one side, for $a = 020$ and $b' = 120$, we will have what we want, namely: $a_1 = 0 \neq 1 = b'_1$ and $g(a) = 0 \neq 2 = g(b')$, and from another one, — unwanted fact, namely: $a_1 = 2 \neq 1 = b'_1$ and $g(a) = 2 = 2 = g(b')$ for $a = 210$ and $b' = 120$, etc.

6. Test for automaton cryptanalytical invertibility

Let q and s, α and β, δ and γ be the values from Q, X^*, X^τ respectively and $a_1a_2a_3, b_1b_2b_3 \in M_\varepsilon$ where a_1, a_2, a_3 and b_1, b_2, b_3 are the different values from $\{q, \alpha, \delta\}$ and $\{s, \beta, \gamma\}$ respectively such that if a_k is q, α or δ, then b_k is s, β or γ respectively, $k \in \{1, 2, 3\}$.

Theorem 1. The automaton A is cryptanalytically invertible of a type $(Q_1x_1Q_2x_2Q_3x_3, \upsilon(q, \alpha, \delta))$, that is, there exists a function f such that (2) is true, if and only if for $Q_1x_1Q_2x_2Q_3x_3$ there is an existential vector function ε such that the following formula is true:

$$\forall a_1a_2a_3 \in M_\varepsilon \, \forall b_1b_2b_3 \in M_\varepsilon (\alpha \neq \beta \Rightarrow ((\bar\varphi(\alpha\delta, q), \upsilon(q, \alpha, \delta)) \neq (\bar\varphi(\beta\gamma, s), \upsilon(s, \beta, \gamma)))). \quad (5)$$

Proof. The proposition under proof is a particular case of Lemma 1. ■

The theorem is the base for deciding by the following exhaustive search method if an automaton A is cryptanalytically invertible of a given type $(Q_1x_1Q_2x_2Q_3x_3, \upsilon(q, \alpha, \delta))$ or not:

1) For every possible existential vector function ε of the $ID = Q_1x_1Q_2x_2Q_3x_3$, generate the existential domain M_ε.
2) Apply Theorem 1 to M_ε, that is, verify whether (5) is true.
3) If for some ε, (5) is true for M_ε, the automaton A is cryptanalytically invertible of the type $(Q_1x_1Q_2x_2Q_3x_3, \upsilon(q, \alpha, \delta))$. Otherwise, that is, if for all existential functions ε of the ID, the condition (5) is false for M_ε, then the automaton A is not cryptanalytically invertible of this type.

7. Decision methods for function cryptanalytical invertibility problem

7.1. Exhaustive search

1) For every possible existential function $\varepsilon = \varepsilon_{j_1} \dots \varepsilon_{j_s}$ where $\varepsilon_k : D_1 \times \dots \times D_{k-1} \to D_k$, $k \in \{j_1, \dots j_s\}$, generate the existential domain M_ε.
2) Apply Lemma 1 to M_ε, that is, verify whether (4) is true.
3) If for some ε, (4) is true for M_ε, the invertibility problem under consideration is positively solvable. Otherwise, that is, if for all ε, (4) is false for M_ε, then the invertibility problem has the negative solution.

7.2. Search by collision elimination

A pair (a,b) of words a and b in $D_1 \times \ldots \times D_n$ is called a collision if $a_{k_0} \neq b_{k_0}$ and $g(a) = g(b)$. We call the collision (a,b) a collision in a subset $U \subseteq D_1 \times \ldots \times D_n$ if $a, b \in U$. Also, we say that U has no collisions, or is free of collisions, if for every a and b in U the pair (a,b) isn't a collision. Further, collisions in an existential domain M_ε as depending on ε are called ε-collisions.

Theorem 2. There exists a function f satisfying (3) if and only if for some ε, the existential domain M_ε has no ε-collisions.

Proof. According to lemma 1,
$\exists f((3)) \Leftrightarrow \exists \varepsilon((4)) \Leftrightarrow \forall a, b \in M_\varepsilon\, (a_{k_0} \neq b_{k_0} \Rightarrow g(a) \neq g(b)) \Leftrightarrow \forall a, b \in M_\varepsilon\, (a_{k_0} = b_{k_0} \vee \vee g(a) \neq g(b)) \Leftrightarrow \forall a, b \in M_\varepsilon \neg(a_{k_0} \neq b_{k_0}\ \&\ g(a) = g(b)) \Leftrightarrow$ (M_ε is free of ε-collisions). ■

Corollary 2. A function $g(x_1, \ldots, x_n)$ is invertible of a type $Q_1 x_1 \ldots Q_n x_n$ with respect to a variable x_{k_0}, $k_0 \in \{i_1, \ldots, i_r\}$, if and only if for some ε, the existential domain M_ε is free of ε-collisions.

For a and b in M_ε, we say as well that the pair (a,b) is a non-ε-collision if it is not a ε-collision, that is, if $a_{k_0} = b_{k_0}$ or $g(a) \neq g(b)$. The following operations are introduced in order to eliminate the ε-collisions from an existential domain M_ε and to get a new domain $M_{\varepsilon'}$ without ε'-collisions (if it is possible) or with other ε'-collisions (otherwise), so witnessing that the function g under consideration is respectively invertible or uninvertible of a given type with respect to a given variable.

Let $a = a_1 \ldots a_{j_1} \ldots a_{j_s} \ldots a_n$, $A = a_{j_1} a_{j_2} \ldots a_{j_s}$, $A' = a'_{j_1} a'_{j_2} \ldots a'_{j_s}$, and $b = b_1 \ldots b_{j_1} \ldots b_{j_s} \ldots b_n$, $B = b_{j_1} b_{j_2} \ldots b_{j_s}$, $B' = b'_{j_1} b'_{j_2} \ldots b'_{j_s}$. Define $a' = a_1 \ldots a_{j_1 - 1} a'_{j_1} \ldots a'_{j_s} a_{j_s + 1} \ldots a_n$ and $b' = b_1 \ldots b_{j_1 - 1} b'_{j_1} \ldots b'_{j_s} b_{j_s + 1} \ldots b_n$. We say that a' and b' are obtained by substituting A by A' and B by B', or A' for A and B' for B, and write $a' = a(A' \to A)$ and $b' = b(B' \to B)$ respectively. Now we can transform the ε-collision (a,b) in M_ε to a non-ε'-collision (a',b) in $M_{\varepsilon'}$ where $a' = a'_1 \ldots a'_n = a(A' \to A)$, $A' \neq A$, $g(a') \neq g(a)$, $\varepsilon' = \varepsilon'_{j_1} \ldots \varepsilon'_{j_s}$, and $\varepsilon'_k(a'_1 \ldots a'_{k-1}) = a'_k$ for each $k \in \{j_1, \ldots, j_s\}$. Analogously, ε-collision (a,b) in M_ε can be transformed to a non-ε'-collision (a,b') in $M_{\varepsilon'}$.

So, in the Example 2, we have ε-collisions (a,b) with $a = 121$ and $b \in \{000, 010, 020\}$ and (a,b) with $a = 201$ and $b \in \{101, 110\}$. In the case $D_3 = \{0,1\}$ that we have it seems impossible to eliminate these ε-collisions without creating others. But if we correct this example and allow $D_3 = \{0,1,2\}$ like in the Example 1, we get a possibility to eliminate them at all by taking, for instance, $\varepsilon'_3(12) = \varepsilon'_3(20) = 2$. In this case ε-collisions $(a,b) = = (121, b)$ and $(a,b) = (201, b)$ in M_ε are transformed to non-ε'-collisions $(122, b)$ and $(202, b)$ respectively in $M_{\varepsilon'}$. At the same time we can note that an elimination of a ε-collision by correcting an existential function ε can produce other collisions and complicate the process of recognizing whether there is an existential function ε without collisions in M_ε. Really, in our example we could eliminate the collisions $(121, b)$ for $b \in \{000, 010, 020\}$ by taking $\varepsilon'(12) = 0$ and obtain the new ε'-collisions $(120, b')$ where $b' \in \{210, 221\}$.

Nevertheless, the notion of ε-collision is very important in the cryptanalytical invertibility theory in many ways. It is enough to say that the exhaustive method above remains strong after changing the need of true condition (4) in it by the request for collision absence (Corollary 2). The requirements of collisions lack in M_ε follow from the need to have a recovering function f (Theorem 2) or invertibility property of g (Corollary 2).

7.3. Search by forward and back tracking

Further, we believe that on any finite set M under consideration a linear ordering relation $\leqslant$ (not greater than) is supposed to be given, and for any $a, b \in M$ we write $a < b$ (a less than b) if $a \leqslant b$ and $a \neq b$. This relation extends to Cartesian products of linearly ordered sets, for instance, as lexicographical ordering in the following way: $a_1 a_2 \ldots a_n < < b_1 b_2 \ldots b_n \Leftrightarrow a_i < b_i$ where i is determined from the conditions $a_1 = b_1, \ldots, a_{i-1} = b_{i-1}$, $a_i < b_i$ and $i \in \{1, 2, \ldots, n\}$, that is, i is the least number in $\{1, 2, \ldots, n\}$ such that $a_i \neq b_i$ and $a_i < b_i$.

Now we introduce some additional notation and notions, namely $I = \{i_1, \ldots, i_r\}$, $i_1 < < \ldots < i_r$, $J = \{j_1, \ldots, j_s\}$, $j_1 < \ldots < j_s$, $n = r + s$, $I \cap J = \varnothing$, $I \cup J = \{1, 2, \ldots, n\}$, $D(I) = \{d_1, \ldots, d_{m_r}\} = D_{i_1} \times \ldots \times D_{i_r}$, $D(J) = \{e_1, \ldots, e_{m_s}\} = D_{j_1} \times \ldots \times D_{j_s}$, $u_{ij} = = d_i \otimes e_j = a_1 a_2 \ldots a_n$ where $a_{i_1} \ldots a_{i_r} = d_i$ and $a_{j_1} \ldots a_{j_s} = e_j$, $i = 1, \ldots, m_r$, $j = 1, \ldots, m_s$, $m = m_r \cdot m_s$, $\{u_1, u_2, \ldots, u_m\} = \{u_{ij} : i = 1, \ldots, m_r, j = 1, \ldots, m_s\}$, $U_t = \{u_1, \ldots, u_t\}$, $1 \leqslant t \leqslant m$.

So, here we consider each vector $a = a_1 a_2 \ldots a_n \in D_1 \times \ldots \times D_n$ as a blend $d \otimes e$ of a vector $d = a_{i_1} a_{i_2} \ldots a_{i_r} \in D(I)$ and a vector $e = a_{j_1} a_{j_2} \ldots a_{j_s} \in D(J)$ and write $a = d \otimes e$.

For every $a = a_1 \ldots a_n \in D_1 \times \ldots \times D_n$ and $b = b_1 \ldots b_n \in D_1 \times \ldots \times D_n$ we say that a and b are equivalent if for each $k \in J$ we have $a_1 \ldots a_{k-1} = b_1 \ldots b_{k-1} \Rightarrow a_k = b_k$. It is clear that this notion here comes from the functionality of the coordinates ε_k of the existential vector function ε. When $a_1 \ldots a_{k-1} = b_1 \ldots b_{k-1}$ and $a_k \neq b_k$ for some $k \in J$, we call the pair (a, b) inequality, and the replacement in a and b the elements a_k and b_k by one and the same element from $D(J)$ is called an inequality elimination. We also say that a subset $U \subseteq D_1 \times \ldots \times D_n$, particularly M_ε, is an equivalence class if all the elements in it are equivalent each other. It is not difficult to see that any such subset is quite simply transformed into an equivalence class by applying, possibly repeatedly, the inequality elimination to pairs of elements in it.

Here in reality, we consider the problem to determine an existential function $\varepsilon : D(I) \to \to D(J)$, that is, for every $d \in D(I)$ to choose an element $\varepsilon(d) \in D(J)$ so that the set $U_\varepsilon = \{d \otimes \varepsilon(d) : d \in D(I)\}$ is namely an equivalence class without collisions (further shortly called ECwC) or to show that such a function ε doesn't exist. The first outcome means that the function $g(x_1, \ldots, x_n)$ is invertible of a given type $Q_1 x_1 \ldots Q_n x_n$ with respect to a given variable x_{k_0}, the second one – that g isn't invertible of this type. The correctness of this decision of the problem is provided by a correct searching an ECwC U_ε with the help of so called forwardtracking (FT) and backtracking (BT) operations correctly defined below and used on the space $D(I) \times D(J)$.

FT: given ECwC $U_t = \{u_1, \ldots, u_t\}$, $1 \leqslant t < m_r$; take $e \in D(J)$ and $u_{t+1} = d_{t+1} \otimes e$ so that u_{t+1} is equivalent to each of $u_1, \ldots, u_t$ and is not in collision with any of them; define $\mathrm{FT}(U_t) = U_{t+1} = \{u_1, \ldots, u_t, u_{t+1}\}$. It is clear that if such an e exists, then FT transforms ECwC U_t into ECwC U_{t+1}. Otherwise, the forwardtracking from ECwC U_t into ECwC U_{t+1} is impossible and backtracking from U_t can be accomplished according to the following general or particular definitions.

BT (general): given ECwC $U_t = \{u_1, \ldots, u_t\}$, $t \geqslant 1$, $d_{t+1} \in D(I)$ and for each $j \in J$ there is $t_j \in \{1, \ldots, t\}$ such that $d_{t+1} \otimes e_j$ isn't equivalent to $u_{t_j} = d_{t_j} \otimes e_{t_j}$ or is in a collision with it. This means that given U_t is impossible to transform by FT into U_{t+1} with given $d_{t+1} \in D(I)$ and any $e_j \in D(J)$ in $u_{t+1} = d_{t+1} \otimes e_j$. In application to these data the backtracking generally consists in taking a specific e from $D(J)$ for $u_{t+1} = d_{t+1} \otimes e$ as well as some $j \in J$ and replacing in U_t the points $u_{t_j} = d_{t_j} \otimes e_{t_j}$ by some other ones $u'_{t_j} = d_{t_j} \otimes e'_{t_j}$ which are equivalent to u_{t+1}, to each other u'_{t_j} and to the rest of U_t and aren't in collision with them. The set $U_{t+1} = U'_t \cup \{u_{t+1}\}$, where u_{t+1} and U'_t are obtained in the described way in the place of d_{t+1} and U_t respectively, is defined as a result of the backtracking from U_t and d_{t+1}, namely: $U_{t+1} = \mathrm{BT}(U_t \cup \{d_{t+1}\})$.

For instance, in Example 1 let $t = 3$, $U_t = \{u_{t-2}, u_{t-1}, u_t\}$, $u_{t-2} = 010$, $u_{t-1} = 101$, $u_t = 120$, $d_{t+1} = 20$, $D(J) = \{e_1, e_2, e_3\}$, $e_1 = 0, e_2 = 1, e_3 = 2$. We can see that every possible value u_{t+1} is in collision with some $u_{t_j} \in \{u_{t-2}, u_{t-1}, u_t\}$. Indeed, if $u_{t+1} = d_{t+1} \otimes e_1$, then $u_{t+1} = 200$ and is in collision with $010 = u_{t-2}$; in analogous way, we can show that if $u_{t+1} = d_{t+1} \otimes e_2 = 201$, then it is in collision with $101 = u_{t-1}$, and if $u_{t+1} = d_{t+1} \otimes e_3 = 202$, then it is in collision with $120 = u_t$. At the same time the set U_t itself is an ECwC, that is, all the points u_{t-2}, u_{t-1}, u_t in U_t are equivalent each other and there are no collisions between them. The aim of backtracking operation is to attach a next data d_{t+1} to a given ECwC U_t as a component of its next member $u_{t+1} = d_{t+1} \otimes e$ admitting the choice of any value e from $D(J)$ as the existential component in u_{t+1} and any correction of existential components in other members of U_t with the only condition — preserving the ECwC properties of the set under backtracking. The choice of the component e in u_{t+1} and the correction of the existential components in members of U_t aren't one-valued and are possible in many different ways. In our instance we have taken $e = 2$ and 0, 1, 2 as existential values in u_{t-2}, u_{t-1}, u_t respectively. So, $U_{t+1} = \mathrm{BT}(U_t \cup \{d_{t+1}\}) = \{u_{t-2}, u_{t-1}, u'_t, u_{t+1}\} = \{010, 101, 122, 202\}$. It is directly verified that this set is ECwC what is required. The following is a result of another variant of backtracking for the same instance: $U_{t+1} = \{u'_{t-2}, u'_{t-1}, u_t, u_{t+1}\} = \{011, 102, 120, 200\}$.

BT (particular): as in general BT, given ECwC $U_t = \{u_1, \ldots, u_t\}$, $d_{t+1} \in D(I)$ and for each $j \in J$ there is $t_j \in \{1, \ldots, t\}$ such that $d_{t+1} \otimes e_j$ isn't equivalent to $u_{t_j} = d_{t_j} \otimes e_{t_j}$ or is in a collision with it. The particular application of the backtracking to these data consists in taking a free (for the first time) existential value e'_t from $D(J)$ for $u'_t = d_t \otimes e'_t$ and e'_{t+1} from $D(J)$ for $u_{t+1} = d_{t+1} \otimes e'_{t+1}$ so that u'_t is equivalent to each of $u_1, \ldots, u_{t-1}$ and without collisions with them, u_{t+1} is equivalent to each of $u_1, \ldots, u_{t-1}, u'_t$ and without collisions with them. The set $U_{t+1} = \{u_1, \ldots, u_{t-1}, u'_t, u_{t+1}\}$ is defined to be the result of the bachtracking from U_t and d_{t+1}, that is, $U_{t+1} = \mathrm{BT}(U_t \cup \{d_{t+1}\})$.

In case, when such an e'_{t+1} doesn't exist, another free e'_t is chosen in $D(J)$ for $u'_t = d_t \otimes e'_t$. If e'_t doesn't exist for choosing a needed e'_{t+1}, then another free e'_{t-1} is chosen in $D(J)$ for $u'_{t-1} = d_{t-1} \otimes e'_{t-1}$ and so on.

After executing BT and constructing U_{t+1} forwardtracking is tried to be applied to U_{t+1}. The computation ends when the beginning point without free existential values is achieved. The analysed function g is adopted to be invertible of a certain type iff an ECwC of the maximal size m_r is demonstrated by this computation.

The author expresses his thanks to I. A. Pankratova for a valuable observation and important correction of the paper.

REFERENCES

1. *Agibalov G. P.* Cryptanalytic concept of finite automaton invertibility with finite delay. Prikladnaya Diskretnaya Matematika. 2019, no. 44, pp. 34–42.
2. *Rasiowa H.* Introduction to Modern Mathematics. Amsterdam; London, North-Holland Publishing Company; Warszawa, PWN, 1973. 339 p.
3. *Agibalov G. P. and Belyaev V. A.* Tehnologiya Resheniya Kombinatorno-logicheskih Zadach Metodom Sokraschyonnogo Obhoda Dereva Poiska [Technology for Solving Combinatorial-Logical Problems by the Method of Shortened Search Tree Traversal]. Tomsk, TSU Publ., 1981. 126 p. (in Russian)

4. *Christofides H.* Graph Theory. An algorithmic Approach. New York; London; San Francisco, Academic Press, 1975.
5. *Zakrevskij A., Pottosin Yu., and Cheremisinova L.* Combinatorial Algorithms of Discrete Mathematics. Tallinn, TUT Press, 2008. 193 p.