Automata Under Effective Observation

Alexandr Babash [0000-0001-7578-6923]

Plekhanov Russian University of Economics, 36 Stremyanny lane, Moscow, 115998, Russia
babash@yandex.ru

Abstract. A trapdoor cipher is a cipher whose algorithm contains some hidden structure (a trapdoor) providing the existence of a subliminal information channel. In cryptographic practice, there could be situations when a constructed cipher may contain some critical defect (a trapdoor) whose identification can significantly weaken the cryptographic strength of this cipher. In this paper, we propose and analyze one of such defects in terms of automata-theoretic approach. An operation of the cipher with this defect is modeled by a finite automaton under the so-called effective observation. The existence of effective observation for a finite automaton qualitatively reflects the presence of a trapdoor which allows one to determine the information on automaton input words by observations over the corresponding output words. We prove the criterion of finding an automaton under effective observation and specify the classes of automata under effective observation and the classes of automata for which there is no effective observation. Possible applications of the results for protecting ciphers from side channel attacks are formulated.

Keywords: Cryptography, Automata theory, Cipher models, Automata under effective observation.

1 Introduction

The analysis of ciphers in terms of automata theory is now becoming quite common. This approach allows one to formulate and solve cryptographic problems for different cipher classes. In cryptographic practice related to the synthesis (construction) of ciphers, there could be situations when one consciously builds a trapdoor cipher, i.e., a cipher whose algorithm contains some hidden structure (a trapdoor) providing the existence of a subliminal information channel. Except for this case, ciphers are usually built without such forethought defects. However, a constructed cipher may

Proceedings of the 10th International Scientific and Practical Conference named after A. I. Kitov "Information Technologies and Mathematical Methods in Economics and Management (IT&MM-2020)", October 15-16, 2020, Moscow, Russia

CEUR Workshop Proceedings (CEUR-WS.org)

contain some critical defect (a trapdoor) whose identification can significantly weaken the cryptographic strength of this cipher. One of such defects is proposed and analyzed in this paper in terms of automata-theoretic approach. This defect presents in a cipher whose operation is modeled by a finite automaton under the so-called effective observation.

Let $A=(X,S,Y,h,f)$ be a finite automaton with an input alphabet X, a set of states S, an output alphabet Y, a transition function $h:S\times X\to S$, and an output function $f:S\times X\to Y$. Denote by $A(s,P)$ an output word of the automaton A resulting from the initial state s when the input word is $P=x1x2\ldots xk$.

The problem. It is required to find functions Uk, Φk such that $Uk(P)=\Phi k(A(s,P))$ for any pair $(s,P)\in S\times Xk$, $Uk\neq const$, $\Phi k\neq const$. An automaton that satisfies this condition for some k is called an automaton under effective observation.

This paper is organized as follows. In section 2, we justify the need for the setting of the presented problem and list possible areas of cryptographic information security where this problem and the results of its solution are of interest. In Section 3, we introduce the basic notions and notations by which automata-theoretic model of the problem is formulated. In this model, the examples of possible problem formulations are presented. Then we prove the criterion of finding an automaton under effective observation, i.e., the criterion of possible determining the information on key elements of a cipher or the information on a plaintext by the given observable information. This section also contains some previous results on the subject obtained by the author. The main results are in Section 5. It is devoted to finding the classes of automata for which, within the given mathematical model, it is impossible to determine the information on an automaton input word by the corresponding output word. We also consider the question whether this problem is algorithmically solvable in the class of all automata. The paper is concluded in Section 6.

2 Relevance of the problem

Let us clarify in which areas of IT technologies it is useful to apply the description of automata under effective observation and the automata for which there is no effective observation. First of all, we are talking about devices and programs that are modeled by finite automata.

2.1 Computer bugs

An information channel is called covert if it is not specifically designed and was not originally supposed to transfer information in electronic data processing system. A covert channel is called subliminal if it can be used only by the holder of the

corresponding secret information. The idea of subliminal information channels was first introduced in the works of Simmons [1-3], by the example of a covert channel in the digital signature system with a public key. In [4], there is the description of the covert channel performance for the digital signature algorithm (DSA) [3], and the papers [5], [6] describe the covert channel performance for the digital signature algorithm Ong-Schnorr-Shamir.

The device embedded in a computer and modeled by a finite automaton under effective observation provides you with an automatic retrieval of the input device information.

2.2 Trapdoor cipher

A trapdoor cipher is a cipher whose algorithm contains some hidden structure (trapdoor) providing the existence of a subliminal information channel; knowledge of this structure allows one to obtain sensitive information (such as secret keys). Without the trapdoor knowledge, the cipher seems to be secure.

One of the most important types of trapdoors that can be embedded in cryptographic algorithms is the so-called SETUP (Secretly Embedded Trapdoor with Universal Protection) mechanism [5], [6]. The SETUP mechanism modifies the given cryptographic algorithm in a way that allows the developer of the cryptosystem to obtain sensitive user information (often about their secret keys). At the same time, for any observer different from the developer, the performance of the modified algorithm is indistinguishable from the performance of the original algorithm. Such modified cryptosystems are sometimes referred as infected cryptosystems.

In [5-7], it is shown that a number of well-established cryptographic primitives can be modified by including the SETUP mechanism in the body of the relevant program. For example, in [7], there is the description of the scheme for constructing a covert information channel in a block cipher. An embedded SETUP mechanism completely compromises the corresponding cryptosystems in relation to its developer (and only in relation to him). These attacks require a single tampering with a cryptosystem. A particular danger is posed by SETUP mechanisms for smart-cards because its key generation always takes place without user intervention. In the case of block ciphers, one can also employ the leakage channels running during the generation and transmission of the so-called Initial Vectors used in OFB, CFB, and CBC modes.

The main known results for embedding trapdoors in ciphers are obtained for public-key cryptosystems, in particular, for the little-known public-key encryption algorithm based on finite automata and called FAPKC (Finite Automaton Public Key Cryptosystem).

Our proposed method for constructing a trapdoor cipher is different from the known ones. It is based on the choice of the automata-theoretic model for the cipher under effective observation. Its principal features are described in the section 5 «On reconstruction of information on an automaton input word by the corresponding output word».

2.3 Side-Channel Attacks

Side-channel attacks (SCA) are a type of cryptographic attacks which exploit information leaked from side channels [8-17]. Let us list some methods to prevent side-channel attacks: masking; blinding; carrying out the computations whose performance time does not depend on data; refusing procedures that use secret intermediate values or keys for conditional transfers; preventing timing attacks by the alignment of performance time for various operations; balancing power consumption; adding a noise (one of the solutions proposed in [18] against Differential Power Analysis (DPA) with the use of a noise is to add random computations which increase the noise level so that it becomes impossible to determine the shift of DPA spikes); shielding; performing double encryption [19]. The approach proposed in the section 5 allows one to put forward an idea of protecting a given automaton from cryptographic attacks using information leaked from side channels. The idea is to construct a new automaton with a large number of states that has the given automaton as a homomorphic image. In the said section, the homomorphic image of the new automaton is an automaton with one state.

2.4 Ciphers

In the works of Claude Shannon [20], there is the description of the so-called "perfect ciphers". In the modern terminology, they are sometimes referred as theoretically unbreakable ciphers. An example of such a cipher is the random gamma cipher. The input alphabet of such a cipher is the set of indexes $I=\{0,1,...,N-1\}$ which label the letters of the ordered alphabet used for plaintexts. Then the set I^L is the set of all possible plaintexts of length L. At the same time, I^L is the set of keys and the set of ciphertexts. The encryption process of the plaintext $i=i_1i_2\ldots i_L$ by the randomly and equiprobably selected key $\gamma=\gamma_1\gamma_2\ldots\gamma_L$ is described by the equations

$$i_j+\gamma_j=z_j \bmod N,\ j\in\{0,1,\ldots,N-1\},$$

where $z=z_1z_2\ldots z_L$ is a ciphertext.

It is obvious that any plaintext i can be encrypted into the ciphertext z under the appropriate choice of the key γ. Thus the plaintext reconstruction based on the

knowledge of a ciphertext is impossible. But if the cipher is not theoretically unbreakable, the known ciphertext sometimes allows one to reconstruct a subset Ii,z of the set IL that contains the desired plaintext. In this case, it is convenient to talk about the information on the plaintext i extracted from its corresponding ciphertext z. The smaller cardinality of this subset, the more extracted information on the desired plaintext. The subset Ii,z can be defined by its characteristic function F1. In cryptographic practice, there could be situations when the ciphertext z is fully unknown. In the introduced notation, this means we know a subset Zz of the ciphertext set Z that contains the given z, or equivalently the value of the characteristic function F2 of this subset.

The cipher for which the desired plaintext can be uniquely determined by its ciphertext is obviously the worst cipher. Not the best cipher would be the cipher for which information on the desired plaintext can be extracted from the observed data of the ciphertext. The above gives rise to the following novel setting of a cryptographic problem. Is it possible, for a given cipher with a plaintext set W, a ciphertext set Z, a set of keys K, and an encryption equation f(w,k)=z, to find nonconstant functions F1 and F2 that satisfy the condition F1(w)=F2(f(w,k)) for any wϵW, kϵK. The conditions under which such functions do not exist are of value for the synthesis of ciphers.

This problem will be solved below using the automata-theoretic approach. It should be noted that by now the automata-theoretic approach to the analysis and synthesis of ciphers has become the natural and traditional direction of cryptanalysis. This is because many encryption systems are mathematically modeled by finite automata. For example, one can imagine that a plaintext is supplied to an input of an encryption device, while the corresponding ciphertext is generated as an output, and the encryption device sequentially changes its state. Note that for a number of classical ciphers, one often imposes the requirement of reversibility which, in the case of decryption error detection, allows one to return to the previous decrypted text and correct the error. In terms of the automata-theoretic approach, the latter is achieved by the use of the so-called permutation automata. Block ciphers are usually also represented by permutation automata, for which reason it is interesting to solve the problem of extracting information on round keys (on an input automaton sequence) from the pairs of input and the corresponding output units (from the pairs of initial and final automaton states).

3 Basic notions and notations

3.1 The notion of an automaton under observation

We will use the following notation:

A=(X,S,Y,h,f) is a finite automaton with an input alphabet X, a set of states S, an output alphabet Y, a transition function h:S×X→S, and an output function f:S×X→Y;

hx:S→S is a partial transition function corresponding to the input symbol x∈X, hxs=h(s,x);

$h_P s = h_{x_k} \ldots h_{x_2} h_{x_1} s$ is a final state of the automaton A resulting from the initial state s when the input word is P=x1x2…xk;

A(s,P) is an output word of the automaton A resulting from the initial state s when the input word is P.

The cardinality of a set M is denoted by |M|.

For a finite automaton A=(X,S,Y,h,f) and a natural number k, denote by Rk, Tk some finite sets. Consider the surjective mappings

$$H_k: S \times X^k \to T_k,$$

$$\Pi_k: S \times X^k \to R_k.$$

Suppose that the automaton A with initial states s from S receives the input words P∈Xk. The pairs (s,P)∈S×Xk for which the automaton operates are unknown. However, for each pair (s,P)∈S×Xk, we know the element t=Hk(s,P) called an observation element.

The problem. Identify the information on the element r=Πk(s,P), where r is the value of the desired operation parameter of the automaton A with initial state s and input word P. By the information on the unknown element r, we mean the identification of a proper subset R` of the set Rk that contains r. Conventionally speaking, we suppose that some function Uk on Rk is given and its value Uk(r)=j defines the subset R`={r`: r`∈Rk, Uk(r`)=j} (lumped state). It is also supposed that j (or the subset R`) can be defined by the known observation element t=Hk(s,P). Formally, we mean that there is a function Φk whose value Φk(t) defines j. Thus we have (Fig.1):

k is a natural number indicating the length of automaton input words;

$\Pi_k: S \times X^k \to R_k$ is an objective function;

$r=\Pi_k(s,P)$ is the value of the parameter under study for the triple (k,A,(s,P)), $(s,P) \in S \times X^k$;

$H_k: S \times X^k \to T_k$ is an observation function;

$t=H_k(s,P)$ is an observation element for the triple $(k,A,(s,P))$, $(s,P) \in S \times X^k$;

Φ_k is an observation information function;

U_k is an objective information function.

The family of introduced objects $(k,Tk,Rk,Hk,\Pi k)$ is called the observation over the automaton A.

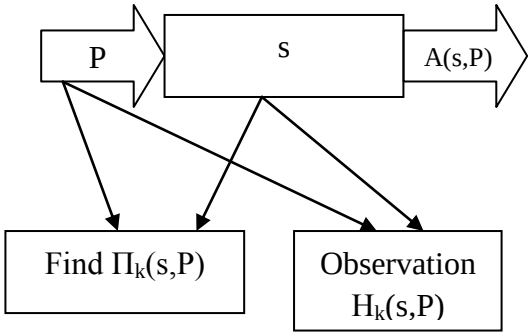

Fig.1. The automaton under effective observation $U_k(\Pi_k(s,P))=\Phi_k(H_k(s,P))$.

Definition 1. An automaton $A=(X,S,Y,h,f)$ is said to be under the effective observation $(k,Tk,Rk,Hk,\Pi k)$ if there exist nonconstant functions Φk on Tk and Uk on Rk such that

$$U_k(\Pi_k(s,P))=\Phi_k(H_k(s,P)) \qquad (1)$$

for any pair $(s,P) \in S \times X^k$.

The property of effective observation can be interpreted as the ability of automatic extraction of the information on the desired parameter $r=\Pi k(s,P)$ from the observation parameter $t=Hk(s,P)$ during the automaton operation process.

Note that if the functions Φk, Uk are constant, equation (1) does not provide any useful information on the value of the parameter r for the observation t.

Definition 2. For the sequence of observations $(k,Tk,Rk,Hk,\Pi k)$, $k \in \{1,2,\ldots\}$ over an automaton A, denote by D(A) the minimum k for which the automaton A is under effective observation. We call D(A) the depth of the automaton observation. If there

is no such k, we set $D(A)=\infty$.

Definition 3. An automaton A is said to be under effective observation if the depth of its sequence of observations is finite.

Definition 4. The functions Φk and Uk for $k=D(A)$ are called the main functions for the automaton A under the observation $(k,Tk,Rk,Hk,\Pi k)$.

3.2 Examples of basic notions

Example 1. Diagnostic experiment with the automaton A [21]. Consider the following observation $(k,Tk,Rk,Hk,\Pi k)$, where $Hk(s,P)=(P,A(s,P))$, Tk is the image of the set $S\times Xk$ under the mapping Hk, $\Pi k(s,P)=s$, and $Rk=S$. The fulfillment of the condition $Uk(\Pi k(s,P))=\Phi k(Hk(s,P))$ for any pair $(s,P)\in S\times Xk$ can be written in the form $Uk(s)=\Phi k(P,A(s,P))$. In this case, the function value Uk(s) is determined by the input and the corresponding output automaton sequences. Suppose further that $Uk(s)=s$ for any $s\in S$. Then the equalities $s=\Phi k(P,A(s,P))$, $(s,P)\in S\times Xk$ mean that each input word $P\in Xk$ is diagnostic for the automaton A [21]. If, instead of all the pairs $(s,P)\in S\times Xk$, we consider only the pairs $(s,P)\in S\times\{P\}$, where $\{P\}$ is a singleton, then the condition $s=\Phi k(P,A(s,P))$, $(s,P)\in S\times\{P\}$ will imply that the word P is a diagnostic sequence for the automaton A.

Homing experiment with the automaton A [21]. The homing experiment formulation in the introduced terms is analogous to that one of the diagnostic experiment.

In [22-24], authors studied the problem of partial definition of an automaton input word by its initial state and the corresponding output word as well as the problem of partial definition of an automaton input word by its final state and the corresponding output word. A more systematic study of these issues was conducted by Sh. Iwen [25], A. A. Kurmit continued this study in his monograph [26].

4 Effective observation over an automaton

4.1 The criterion of finding an automaton A under the effective observation (k,T_k,R_k,H_k,Π_k)

For brevity, we introduce the following notation: $\Leftrightarrow$ means "if and only if"; $\Rightarrow$ means "then"; $\exists$ means "exists"; $\forall$ means "for all"; PQ is a concatenation of words P and Q (in particular, $(P)k$ is a concatenation of k copies of the word P).

Let M be a finite set, σ a reflexive ($m\sigma m\ \forall m\in M$) symmetric binary relation on M, σ^* the transitive closure of σ, and rang σ^* the number of equivalence classes under the binary relation σ^*. The binary relation σ is said to be transitive if rang $\sigma^*=1$.

The functions Hk, Πk induce the partitions of the set $S\times Xk$ defined by the binary

equivalence relations H*=H*k, Π*=Π*k on S×Xk. Namely,

$$(s,P)H^*(s`,P`) \Leftrightarrow H_k(s,P)=H_k(s`,P`),$$

$$(s,P)\Pi^*(s`,P`) \Leftrightarrow \Pi_k(s,P)=\Pi_k(s`,P`).$$

The classes under the equivalence relations H*k, Π*k will be denoted by the letters t and r of the corresponding sets Tk, Rk, i.e, the letter t can also denote Hk-1(t), the preimage of t, and similarly for the letter r. The context will make it clear whether t and r stand for elements or classes. In particular, Tk, Rk will refer, if necessary, to the sets of the classes under the equivalence relations H*k, Π*k, respectively.
Denote by t[s,P] the class comprising (s,P) under the equivalence relation H*k and by r[s,P] the class comprising (s,P) under the equivalence relation Π*k.
Let us introduce the binary equivalence relations Tk/Π*k on Tk and Rk/H*k on Rk. For this purpose, we use the auxiliary binary relations ~ on Tk and ~ on Rk. The same symbol ~ is used for simplicity. The elements t, r of the sets Tk, Rk are simultaneously considered as the classes t[s,P], r[s,P],

$$t\sim t` \Leftrightarrow \exists(s,P)\in t,\ (s`,P`)\in t`,\ r\in R_k:$$

$$(s,P)\in r,\ (s`,P`)\in r.$$

The equivalence relation T_k/Π^*_k is the transitive closure of the binary relation ~, i.e.,

$$t_1\ T_k/\Pi^*_k\ t_L \Leftrightarrow \exists t_2,\ldots,t_{L-1}\in T_k:\ t_1\sim t_2\sim\ldots\sim t_{L-1}\sim t_L.$$

Similarly,

$$r\sim r` \Leftrightarrow \exists(s,P)\in r,\ (s`,P`)\in r`,\ t\in T_k:$$

$$(s,P)\in t,\ (s`,P`)\in t,$$

and R_k/H^*_k is the transitive closure of the binary relation ~ on R_k.

Note that

$$\forall t\in T_k,\ r\in R_k\colon t\sim t,\ r\sim r,\ t\ T_k/\Pi^*_k\ t,\ r\ R_k/H^*_k\ r.$$

Define the binary equivalence relation $H^*k\vee\Pi^*k$ on $S\times Xk$ by the auxiliary binary relation $\approx\approx$ on $S\times Xk$,

$$(s,P)\approx\approx(s`,P`)\Leftrightarrow\exists t\in T_k,\ r\in R_k\colon$$

$$(s,P),(s`,P`)\in t\vee r,$$

i.e., $(s,P),(s`,P`)\in t$ or $(s,P),(s`,P`)\in r$. To specify the relation

$$(s,P)\approx\approx(s`,P`),\ (s,P),(s`,P`)\in t\vee r,$$

we write

$$(s,P)\approx H^*_k\approx(s`,P`)\Leftrightarrow\exists t\in T_k\colon$$

$$(s,P),(s`,P`)\in t$$

and

$$(s,P)\approx\Pi^*_k\approx(s`,P`)\Leftrightarrow\exists r\in R_k\colon$$

$$(s,P),(s`,P`)\in r.$$

Denote by $H^*k\vee\Pi^*k$ the transitive closure of the binary relation $\approx\approx$.
The equivalence relations Tk/Π^*k on Tk and Rk/H^*k on Rk induce the equivalence relations $(Tk/\Pi k)^*$, $(Rk/Hk)^*$ on $S\times Xk$,

$$(s,P)\ (T_k/\Pi_k)^*\ (s`,P`)\Leftrightarrow$$

$$t[s,P]\ T_k/\Pi^*_k\ t[s`,P`],$$

$$(s,P)\ (R_k/H_k)^*\ (s`,P`)\Leftrightarrow$$

$$r[s,P]\ R_k/H^*_k\ r[s`,P`].$$

Denote by $H^*k\vee\Pi^*k[s,P]$, $(Tk/\Pi k)^*[s,P]$, $(Rk/Hk)^*[s,P]$ the classes comprising (s,P) under the equivalence relations $H^*k\vee\Pi^*k$, $(Tk/\Pi k)^*$, $(Rk/Hk)^*$ on $S\times Xk$, respectively.

Theorem 1. For the binary equivalence relations on $S\times Xk$ there holds the equality $H^*k\vee\Pi^*k=(Tk/\Pi k)^*=(Rk/Hk)^*$.

Proof. Let us show that $(Tk/\Pi k)^*=H^*k\vee\Pi^*k$. The equality $(Rk/Hk)^*=H^*k\vee\Pi^*k$ can be proved similarly. Suppose $(s,P)\ (Tk/\Pi k)^*\ (s`,P`)$. Then

$$(s,P)\ (T_k/\Pi_k)^*\ (s`,P`)\Leftrightarrow$$

$$t_1[s,P]\ T_k/\Pi^*_k\ t_L[s`,P`]\Rightarrow$$

$$\exists t_2,\ldots,t_{L-1}\in T_k:$$

$$t_1[s,P]\sim t_2\sim\ldots\sim t_{L-1}\sim t_L[s`,P`]\Rightarrow$$

$$t_j\sim t_{j+1},\ j\in\{1,2,\ldots,L-1\}\Rightarrow$$

$$\exists(s_j,P_j)\in t_j,\ (s`_j,P`_j)\in t_{j+1},\ r_j\in R_k:$$

$$(s_j,P_j),\ (s`_j,P`_j)\in r_j\Rightarrow(s,P)\approx H^*_k\approx(s_1,P_1)\approx\Pi^*_k\approx(s`_1,P`_1)\approx H^*_k\approx(s_2,$$

$$P_2)\approx\Pi^*_k\approx(s`_2,P`_2)\approx\approx\ldots\approx\approx(s`,P`)\Rightarrow$$

$$(s,P)\ H^*_k\vee\Pi^*_k\ (s`,P`).$$

Conversely, suppose $(s1,P1)H^*k\vee\Pi^*k(sL,PL)$. Then

$$(s_1,P_1)\ H^*_k\vee\Pi^*_k\ (s_L,P_L)\Rightarrow\exists(s_2,P_2),\ldots,(s_{L-1},P_{L-1}),\ \varepsilon(1),\varepsilon(2),\ldots,\varepsilon(L-1)\in\{H^*_k,\ \Pi^*_k\}:$$

$$(s_1,P_1)\approx\varepsilon(1)\approx(s_2,P_2)\approx\varepsilon(2)\approx,\ldots,\ (s_{L-1},P_{L-1})\approx\varepsilon(L-1)\approx(s_L,P_L).$$

In the above chain of binary relations, choose the chain of minimum length L. In this case, $\varepsilon(j)\neq\varepsilon(j+1)$ for any $j\in\{1,2,\ldots,L-2\}$ and we have

$$t[s_j,P_j]=t[s_{j+1},P_{j+1}]\text{ if }\varepsilon(j)=H^*_k,$$

$$t[s_j,P_j]\sim t[s_{j+1},P_{j+1}] \text{ if } \varepsilon(j)=\Pi^*_k.$$

Hence,

$$\exists t_2,\ldots,t_{L`}: t[s_1,P_1]\sim t_2,\sim\ldots\sim t_{L`}\sim t[s_L,P_L] \Rightarrow$$

$$t[s_1,P_1]\ T_k/\Pi^*_k\ t[s_L,P_L] \Rightarrow$$

$$(s_1,P_1)\ (T_k/\Pi_k)^*\ (s_L,P_L).$$

□

Corollary 1. For the number of classes under the equivalence relations $H^*k\vee\Pi^*k$ on $S\times Xk$, Tk/Π^*k on Tk, and Rk/H^*k on Rk, it follows that rang $H^*k\vee\Pi^*k$=rang Tk/Π^*k =rang Rk/H^*k.

Proof. In view of Theorem 1, it suffices to show that

$$\text{rang } T_k/\Pi^*_k=\text{rang } (T_k/\Pi_k)^*.$$

By definition,

$$(s,P)\ (T_k/\Pi_k)^*\ (s`,P`) \Leftrightarrow$$

$$t[s,P]\ T_k/\Pi^*_k\ t[s`,P`],$$

from whence it follows that

$$(T_k/\Pi_k)^*[s,P]=\cup t_j[s_j,P_j]$$

for any $(s,P)\in S\times Xk$, where the union is taken over all the classes $tj\in Tk/\Pi^*k[t[s,P]]$. The required statement follows directly from the above equation.

□

Theorem 2. The automaton A=(X,S,Y,h,f) is under the effective observation (k,Tk,Rk,Hk,Πk) if and only if rang $H^*k\vee\Pi^*k\geq 2$. Moreover, the values of the functions ΦkHk, UkΠk on $S\times Xk$ are the same and constant on the classes $H^*k\vee\Pi^*k$, while the values of the functions Φk, Uk are constant on the classes Tk/Π^*k and Rk/H^*k, respectively.

Proof. Suppose that some functions Φk, Uk satisfy the following:

$$\Phi_k(H_k(s,P))=U_k(\Pi_k(s,P))$$

for any $(s,P)\in S\times Xk$.

Consider two elements (s1,P1), (sL,PL) from the same class under the equivalence relation $H^*k\vee\Pi^*k$. Then

$$\exists\varepsilon(1), \varepsilon(2), \ldots, \varepsilon(L-1)\in\{H^*_k, \Pi^*_k\}, (s_2,P_2),\ldots,(s_{L-1},P_{L-1})\in S\times X^k:$$

$$(s_1,P_1)\approx\varepsilon(1)\approx(s_2,P_2)\approx\varepsilon(2)\approx\ldots\approx\varepsilon(L-1)\approx(s_L,P_L).$$

For $\varepsilon(j)=H^*k$, we obtain $\Phi k(Hk(sj,Pj))=\Phi k(Hk(sj+1,Pj+1))$ and hence

$$U_k(\Pi_k(s_j,P_j))=U_k(\Pi_k(s_{j+1},P_{j+1})).$$

For $\varepsilon(j)=\Pi^*k$, we obtain $Uk(\Pi k(sj,Pj))=Uk(\Pi k(sj+1,Pj+1))$ and hence

$$\Phi_k(H_k(s_j,P_j))=\Phi_k(H_k(s_{j+1},P_{j+1})).$$

Thus the functions ΦkHk, UkΠk on $S\times Xk$ take the same and constant values on the classes $H^*k\vee\Pi^*k$. By Theorem 1,

$$H^*_k\vee\Pi^*_k=(T_k/\Pi_k)^*=(R_k/H_k)^*,$$

and hence the functions ΦkHk, UkΠk are constant on the classes under the equivalence relations $(Tk/\Pi k)^*$, $(Rk/Hk)^*$.

For any pair $(s,P)\in S\times Xk$,

$$(T_k/\Pi_k)^*[s,P]=\bigcup_{t_j} t_j[s_j,P_j],$$

where the union is taken over all $tj\in Tk/\Pi^*k[t[s,P]]$. Similarly,

$$(R_k/H_k)^*[s,P]=\bigcup_{r_j} r_j[s_j,P_j],$$

where the union is taken over all $rj\in Rk/H^*k[r[s,P]]$. Therefore, Φk and Uk are constant on the classes Tk/Π^*k and Rk/H^*k, respectively. It is clear that Φk and Uk are nonconstant if and only if

$$\text{rang } H^*_k\vee\Pi^*_k=\text{rang } T_k/\Pi^*_k=$$

$$\text{rang } R_k/H^*_k\geq 2.$$

□

Corollary 2. The depth D(A) of the observation over the automaton A for the sequence of observations (k,Tk,Rk,Hk,Πk), k∈{1,2,…} coincides with the minimum k for which rang H*k∨Π*k=rang Tk/Π*k=rang Rk/H*k≥2. If there is no such k, then D(A)=∝.

4.2 Previous results

Reconstruction of information on the first input symbol of an automaton input word by the corresponding initial state and output sequence [27]: Rk=X, Tk=(S×Yk), k∈{1,2,…};

$$\Pi_k(s,x_1x_2\ldots x_k)=x_1,\ H_k(s,x_1,x_2,\ldots,x_k)=(s,A(s,x_1,x_2,\ldots,x_k)).$$

Reconstruction of information on an input word in a permutation automaton by the corresponding initial state and output sequence [27]:.

$$R_к=X^к,\ \Pi_к:X^к\times S\to X^к,\ \Pi_к(x_1,x_2,\ldots,x_к,s)=x_1,x_2,\ldots,x_к,\ T_к=S\times Y^к,\ H_к:X^к\times S\to S\times Y^к,$$

$$H_к(x_1,x_2,\ldots,x_к,s)=(s,A(s,x_1,x_2,\ldots,x_к)).$$

Reconstruction of information on an automaton input word by the corresponding final state and output sequence [27]: Rк=Xк, Пк:Xк×S→Xк, Пк(x(1)x(2)…x(к),s)=(x(1)x(2)…x(к)), Тк=S×Yк, Нк:Xк×S→S×Yк, Нк(x(1)x(2)…x(к),s)=(h(s,x(1)x(2)…x(к)), A(s,x(1)x(2)…x(к))).

Reconstruction of information on an input word in a permutation automaton given initial and final states [28].

For a word P=x(1)x(2)…x(k) in Xk, set hP=hx(k)hx(k-1)…hx(1). Denote by hPs the image of s under hP and by S(L) the set of all subsets of S with cardinality L, L≥1. Define the function Hk,L on Xk×S(L) as follows. For s(L)={sj(1),sj(2),…,sj(L)} in S(L) and P=x(1)x(2)…x(k) in Xk, set

$$H_{k,L}(P,s^{(L)})=(s^{(L)};h_Ps^{(L)})=$$

$$\{(s_{j(1)},h_Ps_{j(1)}),(s_{j(2)},h_Ps_{j(2)}),\ldots,(s_{j(L)},h_Ps_{j(L)})\}.$$

This expression can be considered as a partial permutation on S (L transitions are defined for the permutation hP). Denote by Hk,L(Xk×S(L)) the image of Hk,L.

We say that k-length input words of an automaton A can be approximately reconstructed given L initial and final states if there exist nonconstant functions Φk,L and Uk,L, defined on Hk,L(Xk×S(L)) and Xk, respectively, such that for any P∈Xk and s(L)∈S(L) we have

$$U_{k,L}(P)=\Phi_{k,L}(s^{(L)},h_P s^{(L)}).$$

In what follows, we refer to this fact by saying that the automaton A possesses the (Xk,s,hPs,L)-reconstruction property.

The problem of describing automata with the (Xk,s,hPs,L)-reconstruction property is close to the following problems: experimental designs for automata (see [24], [30]), where an unknown input word is used for testing, with initial and final states being observed as experimental results; local reconstruction of information on an input word (see [29], [30]) given initial and final states; description of information-lossless automata [24]. A number of results concerning the problem under consideration are presented by the author in [31-35].

5 On reconstruction of information on an automaton input word by the corresponding output word

5.1 Setting of the problem

In what follows, T_k is the set of all k-length output words $A(s,P)$ of the automaton A corresponding to the input word $P\in X^k$ and the initial state $s\in S$, $R_k=X^k$, $H_k(s,P)=A(s,P)$, $\Pi_k(s,P)=P$.

Such an observation depends on the parameter k. From a cryptographic point of view, this naturally gives rise to the question whether there exists a natural k for which the given automaton A=(X,S,Y,h,f) is under the observation (k,T_k,R_k,H_k,Π_k). In the case of positive answer to this question, one naturally comes to the further problem of the upper estimate for such a minimum k. If there is no such k, then the cipher that is modeled by such an automaton is particularly valuable.

In order to describe the classes of automata under the effective observation $(k,T_k,R_k,H_k,\Pi_k)=(R_k=X^k,\Pi_k(s,P)=P,\ H_k(s,P)=A(s,P))$ and the automata for which there is no effective observation, we introduce the following notions and notation.

Denote by X* the set of all finite-length words over the alphabet X and by X^{∞} the set of all infinite words over the alphabet X. For the concatenation of words $P \in X^n$, $P'' \in X^m$, we write PP", where P is an initial subword of PP". If n=0, then any word from X^n is considered to be empty and PP"=P". For the subwords of $P=x_1,x_2,...$, we use the notation:

$P^1=P$; $P^n=x_nx_{n+1}...\ P]_m=x_1x_2...x_m\ P^n]_m=x_nx_{n+1}...,x_m \quad n \le m$; $(P)^k=PP\ldots P$ k times.

Define a number of binary relations on the sets X^{∞} and X^L, $L \in \{1,2,\ldots\}$.

The relation σ_L on X^L

$$P\sigma_L P'' \Leftrightarrow \exists s,s'' \in S: A(s,P)=A(s'',P'').$$

The relation σ_∞ on X^{∞}

$$P\sigma_\infty P'' \Leftrightarrow \exists s,s'' \in S: A(s,P)=A(s'',P'').$$

The relation $\sigma_\infty]_L$ on X^L induced by the binary relation σ_∞

$$P_1\sigma_\infty]_L P_2 \Leftrightarrow \exists P,P'' \in X^{\infty}: P\sigma_\infty P'',\ P]_L=P_1,\ P'']_L=P_2.$$

Denote by $(\sigma_\infty)^*$, $(\sigma_\infty]_L)^*$, $(\sigma_L)^*$ the transitive closures of the binary relations σ_∞, $\sigma_\infty]_L$, σ_L. For the binary relation $\tau \in \{(\sigma_\infty), (\sigma_\infty]_L), (\sigma_L)\}$, denote by rang τ the number of equivalence classes under the binary equivalence relation $\tau^*$and by $t(\tau)$ the minimum L for which

$$\text{rang}(\sigma_L)^*>1 \qquad (\text{rang}(\sigma_\infty]_L)^*>1).$$

Otherwise, set $t((\sigma_L)^*)=\infty$ $(t((\sigma_\infty]_L)^*)=\infty)$. The introduced parameters are

called the degrees of the transitive binary relations σ_L, $\sigma_\infty]_L$, respectively.

The next proposition follows from Theorem 2.

Proposition 1. *The automaton A is under the effective observation (k, $R_k=X^k$, $\Pi_k(s,P)=P$, $H_k(s,P)=A(s,P)$) if and only if rang(σ_k)>1.*

5.2 The class of automata without effective observation

The class of automata with the loss of information (B1 type) and without effective observation. Now let us turn to the description of the automaton classes for which the observation (k, $R_k=X^k$, $\Pi_k(s,P)=P$, $H_k(s,P)=A(s,P)$) is not effective for any k.

Definition 5. An automaton A=(X,S,Y,h,f) is said to be an automaton with the loss of information B1 if for $L \geq \frac{|S|(|S|-1)}{2}+1$ there exists a binary relation ε on the set X with the following properties:

1) For each pair (x,x") from ε there exist P,P" from X^L and s∈S such that A(s,xP)=A(s,x"P");

2) The binary relation ε is transitive on the set X.

Theorem 3. *If a permutation automaton A is an automaton with the loss of information (B1 type), then its observation (k, $R_k=X^k$, $\Pi_k(s,P)=P$, $H_k(s,P)=A(s,P)$) is not effective for any k.*

Proof. Suppose the conditions of the theorem are satisfied. Let us first prove that the theorem conditions imply the following: for any pair (x,x")∈ε and any

word $P \in X^*$ there exist Q, $Q'' \in X^{\infty}$ such that $PxQ\sigma_{\infty}Px''Q''$. Indeed, suppose there is a pair $(x,x'') \in \varepsilon$. Then there exist automaton input words $P=p_1p_2\ldots p_L$, $P''=p''_1p''_2\ldots p''_L$ and an initial state $s \in S$ such that $A(s,xP)=A(s,x''P'')$. In the transition graph of the automaton, the words P, P'' are represented as the paths from the initial state $s \in S$ passing through the states $s, h_{p_1}s, h_{p_2}h_{p_1}s, \ldots, h_{p_L}\ldots h_{p_2}h_{p_1}s$ and $s, h_{p''_1}s, h_{p''_2}h_{p''_1}s, \ldots, h_{p''_L}\ldots h_{p''_2}h_{p''_1}s$, respectively. The two following cases are possible: 1) there exists a pair of states $(h_{p_j}\ldots h_{p_2}h_{p_1}s, h_{p''_j}\ldots h_{p''_2}h_{p''_1}s)$ such that $h_{p_j}\ldots h_{p_2}h_{p_1}s = h_{p''_j}\ldots h_{p''_2}h_{p''_1}s$, 2) the states in each pair of states $(h_{p_j}\ldots h_{p_2}h_{p_1}s, h_{p''_j}\ldots h_{p''_2}h_{p''_1}s)$, $j \in \{1,2,\ldots L\}$ are different. By standard methods of graph theory, it can be shown that in each of the cases 1), 2)

$$Px\sigma_{\infty}]_{k+1}Px''$$

for any $k \in \{0,1,\ldots\}$, $P \in X^k$, and $(x,x'') \in \varepsilon$.

In particular, $x\sigma_{\infty}]_1x''$. Hence, $\mathrm{rang}(\sigma_{\infty}]_1)^*=1$. Thus we obtain $\mathrm{rang}(\sigma_1)^*=1$ on X^1. Assume that for some K, $\mathrm{rang}(\sigma_{\infty}]_K)^*=1$. Then let us prove that $\mathrm{rang}(\sigma_{\infty}]_{K+1})^*=1$. By the assumption, it follows that there exists a chain of binary relations

$$P_1\sigma_{\infty}]_KP_2\sigma_{\infty}]_KP_3\ldots\sigma_{\infty}]_KP_N,\ N \geq |X|^K$$

that contains all the words from X^K.

From the definition of the binary relation $\sigma_{\infty}]_K$ on X^K it follows that

$$\exists \alpha_j,\beta_j \in X: P_j \alpha_j \sigma_\infty]_{K+1} P_{j+1}\beta_j,\ j \in \{1,2,\ldots,N\}.$$

To complete the proof of transitivity of the binary relation $\sigma_\infty]_K$ on X^K for any K, we introduce the binary relation $\varepsilon(Q)$ on the set of words $\{Qx: x \in X\}$, where Q is an arbitrary finite word of length K over the alphabet X. Let $\varepsilon(Q)$ be a subset of the set $(X^K)^2$ that consists of all the pairs (Qx,Qx"), where $(x,x") \in \varepsilon$.

Lemma 1. *The binary relation $\varepsilon(Q)$ on the set of words $\{Qx: x \in X\}$, where $Q \in X^K$, is transitive.*

Proof. For the automaton A, fix all the triples (s,x,x") with the property: for each pair (x,x") from ε there exist P, P" from X^∞ and $s \in S$ such that A(s,xP)=A(s,x"P"). Since the automaton A is a permutation automaton, it follows that for any triple (s,x,x") and any automaton input word Q there exists a state s_Q such that $h_Q s_Q = s$. It is obvious that $A(s_Q,Qx)=A(s_Q,Qx")$. For this reason, and by virtue of transitivity of the binary relation ε, the binary relation $\varepsilon(Q)$ is transitive. Thus Lemma 1 is proved.

□

Since the set of words X^{K+1} is divided into disjoint subsets $\varepsilon(Q)$, $Q \in X^K$, on each of which the binary relation ε is transitive, and the subsets are related by (2), the induction step is carried out. We now note that the transitivity of the binary relation $\sigma_\infty]_{K+1}$ on X^{K+1} implies the transitivity of the binary relation σ_{K+1} on X^{K+1}.

□

The class of automata with the loss of information (B2 type) and without

effective observation.

Definition 6. An automaton A=(X,S,Y,h,f) is said to be an automaton with the loss of information (B2 type) if for $L \geq \frac{|S|(|S|-1)}{2} + |S|$ there exists a binary relation ε on the set X with the following properties:

1) For each pair (x',x") from ε there exist P',P" from X^L and s, s"∈S such that

A(s,P'x)=A(s",P"x"), $h_{P'x'}s = h_{P"x"}s"$.

2) The binary relation ε is transitive on the set X.

Theorem 4. *If an automaton A=(X,S,Y,h,f) is an automaton with the loss of information (B2 type), then its observation (k, $R_k=X^k$, $\Pi_k(s,P)=P$, $H_k(s,P)=A(s,P)$) is not effective for any k.*

Proof. Suppose that for L=0, X^L consists of an empty word. We will use the previously introduced notation σ_L of the binary relation on X^L. Denote by $\sigma(L)$ the new binary relation on X^L,

P' $\sigma(L)$ P" ⇔ ∃ words P_1', $P_1"$∈ X^{N_1}, P_2', $P_2"$∈ X^{N_2}, where $N_1=N_1(P',P")\geq 1$, $N_2=N_2(P',P")\geq 0$: ∀k≥1 it follows that $(P_1')^k P_2' P' \sigma_{kN_1+N_2+L}$ $(P_1")^k P_2" P"$.

Lemma 2. *If A=(X,S,Y,h,f) is an automaton with the loss of information (B2 type), then*

$\forall x', x'' \in \varepsilon,\ P \in X^L,\ L \in \{0,1,2,\ldots\}:\ x' P\, \sigma(L+1)\, x'' P.$

$A(s,P'x)=A(s'',P''x''),\ \ h_{P'x'}s = h_{P''x''}s''.$

2) The binary relation ε is transitive on the set X.

6 Discussion

We note the universality of the approach to the study of useful (key) information of a finite automaton based on its possible observation. In this connection, the range of topical problems of finding information about the automaton of interest to the researcher can be significantly expanded. The question of existence of an algorithm recognizing the effective observation (k, Rk=Xk, Πk(s,P)=P, Hk(s,P)=A(s,P)) in the class of all automata remains opened.

7 Conclusion

In this paper, we have introduced some novel notions of the automata-theoretic approach: observation over an automaton and an automaton under effective observation. The latter qualitatively reflects the presence of an automaton trapdoor for determining the information on an automaton input word by the observable information. For a given observation, we have proved the criterion of finding an automaton under effective observation. Also we have specified the classes of automata under effective observation and the classes of automata for which there is no effective observation.

References

1. Simmons G J. The subliminal channel and digital signatures. EUROCRYPT'84, Lect. Notes Comput. Sci., 1985, v.209: 51-57.
2. Simmons G J. A secure subliminal channel (?). CRYPTO'85, Lect. Notes Comput. Sci., 1986, v.218: 33-41.
3. Simmons G J. Subliminal communication is easy using the DSA. EUROCRYPT'93, Lect. Notes Comput. Sci., 1994, v.765: 218-232.

4. Digital Signature Standard (DSS). Federal Information Processing Standard (FIPS) Publication 186, National Institute of Standards and Technology (NIST), US Department of Commerce, Washington D.C., December, 1998.
5. Young A, Yung M. The Dark Side of Black-Box Cryptography. CRYPTO'96, Lect. Notes Comput. Sci., 1997, v.1109: 89-103.
6. Young A, Yung M. Kleptography: using Cryptography against Cryptography. EUROCRYPT'97, Lect. Notes Comput. Sci., 1998, v.1233: 62-74.
7. Young A, Yung M. Monkey. Black-Box Symmetric Ciphers Designed for monopolizing keys. FSE'98, Lect. Notes Comput. Sci., 1998, v.1372: 122-133.
8. YongBin Zhou, DengGuo Feng. Side-Channel Attacks: Ten Years After Its Publication and the Impacts on Cryptographic Module Security Testing.
9. P. Kocher. Timing attacks on implementations of Diffie-Hellmann, RSA, DSS, and other systems. CRYPTO'96, LNCS, 1996, v.1109: 104-113.
10. J. Black, H. Urtubia. Side-channel attacks on symmetric encryption schemes: the case for authenticated encryption. In Proc of 11th USENIX Security Symposium, 2002, pp.327-338.
11. D. Boneh, R.A. DeMillo, R.J. Lipton. On the importance of checking cryptographic protocols for faults. EUROCRYPT'97, LNCS, 1997, v.1233: 37-51.
12. K. Gandolfi, C. Mourte, F. Olivier. Electromagnetic Analysis: Concrete Results. CHES 2001, LNCS, v.2162: 251-261.
13. J. J. Quisquater, D. Samyde. Electromagnetic analysis (EMA): measures and counter-measures for smart cards. E-smart 2001, LNCS, v.2140: 200–210.
14. M. G. Kuhn, R. J. Anderson. Soft tempest: hidden data transmission using electromagnetic emanations. Information Hiding 1998, LNCS, v.1525: 124-142.
15. M. Kuhn. Optical Time-Domain Eavesdropping Risks of CRT Displays. Proc of the 2002 Symposium on Security and Privacy, pp.3-18.
16. A. Shamir, E.Tramer. Acoustic cryptanalysis: on nosy people and noisy machines. In Eurocrypt 2004 rump session.
17. Y. Tsunoo, T. Saito, T. Suzaki, M. Shigeri, H. Miyauchi.Cryptanalysis of DES Implemented on Computers with Cache. CHES 2003, LNCS, v.2779: 62–76.
18. T. S. Messerges, E. A. Dabbish, R. H. Sloan, Examining smart-card security under the threat of power analysis attacks. IEEE Trans. Computers, 2002, 51(5): 541-552.
19. E. Biham, A. Shamir. Differential fault analysis of secret key cryptosystems. CRYPTO'97, LNCS, 1997, v.1294: 513-525.
20. C. Shannon. Raboti po teorii informacii i kibernetike (Works on information theory and cybernetics). Inostran. Lit., Moscow, 1963.
21. A. Gill. Introduction to the Theory of Finite-State Machines, McGraw-Hill Book Co., New York, 1962.

22. Huffman D A. Canonical forms for information-lossless finite-state logical machines. IRE Trans. Circuit Theory, 6, Spec. Suppl., 1959, 41-59.
23. Huffman D A. Notes on information-lossless finite-state automata. Nuovo cimento, 13, Suppl. 2, 1959, 397-405.
24. Even Sh. On information-lossless automata of finite order. IEEE Trans. Electronic Comput., 14, 1965, 4, 561-569.
25. Even Sh. Ob avtomatah konechnogo poryadka bez poteri informacii. Trudi mejdunarodnogo simpoziuma po teorii releinih ustroistv i konechnih avtomatov (On information-lossless automata of finite order. Proceedings of the International Symposium on the theory of relay devices and finite automata.). Moscow: Nauka, 1965, 269-279.
26. Kurmit A. A. Avtomati bez poteri informacii konechnogo poryadka (On information-lossless automata of finite order). Riga: Zinatne, 1972, 264.
27. Babash A. V. Ob eksperimentah po raspoznavaniyu informacii o vhodnom slove avtomata (On experiments on recognition of information on an automaton input word). Proceeding on Discrete Mathematics, v.8., Moscow: FIZMATLIT, 2004, pp.7-24.
28. Babash A. V. On Reconstruction of Information on an Input Word in a Medvedev Permutation Automaton Given Initial and Final States. Problems of Information Transmission, 2007, Vol. 43, No. 2, pp.132–142.
29. Bogomolov A.M., Barashko A.S., and Grunskii I.S., Eksperimenty s avtomatami (Experiments with Automata), Kiev: Naukova Dumka, 1973.
30. Tverdokhlebov V.A., Logicheskie eksperimenty s avtomatami (Logical Experiments with Automata), Saratov: Izd. Sarat. Univ., 1988.
31. Babash A. V. Neotlichimost sostoyanii konechnogo avtomata otnositelno funkcii, zadannoi na ego vhodnih i vihodnih slovah (Indistinguishability of finite automaton states for the function defined on its input and output words). Obozr. Prikl. Prom. Mat., 2001, vol. 8, no. 1, pp.94–95.
32. Babash A. V. On Some Invariants of a Finite Automaton. Obozr. Prikl. Prom. Mat., 2000, vol. 7, no. 1, pp.86–87.
33. Babash A. V., Local Reconstruction of Input Words of Automata Given the Initial and Final States. In Proc. 2nd All-Russian Symp. on Applied and Industrial Mathematics, Samara, 2001, pp.93–94.
34. Glukhov M. M., On Numerical Parameters Associated with the Definition of Finite Groups by Systems of Generating Elements. In Trudy po diskretnoi matematike (Proceedings in Discrete Mathematics), vol. 1, Moscow: TVP, 1997, pp.43–66.
35. Krapeˇz A., On a Generalization of Fermat's Theorem in the Theory of Groups, Publ. Inst. Math.(Beograd) (N.S.), 1974, vol. 17(31): 77–81.