ESTIMATION OF SOME ENCRYPTION FUNCTIONS IMPLEMENTED INTO SMART CARDS

H. Groscot

76, rue A. Briand, 93220, Gagny, France .

Abstract .

We study a family of encryption functions , wich is particularly adapted for the situations that arise in smart cards .

Probabilistic arguments show us that "big key" is not synonymous of "good security" for these functions .

We think that the security of such functions has to rely on other criteria .

T. Beth, N. Cot, and I. Ingemarsson (Eds.): Advances in Cryptology - EUROCRYPT '84, LNCS 209, pp. 470-479, 1985.

I. GENERALITIES

1. About protocols using smart cards .

The interest of smart cards not only lies in their intelligence but their security performances concern us as well . These allow the implementation of encryption functions , in order to perform authentication procedures, as those needed in long-distances payments or in checking acces control to buildings .

In this paper, we first show on a small example that smart cards allow us to design cryptographic protocols . In this example, it is assumed that the cards are issued by an organisation on business purpose and may accomplish long-distances operations .

They contain an encryption function γ depending on a secret key S, wich is the same in all the cards .

Before each transaction with a customer at a sail's point, one has to be sure that the customer's card is a valid one, therefore contains γ and S. So one connects the sail's point to a central by means of a telecommunication network. This central has a valid card and a pseudo random generator at his disposal.

After the insertion of the card in the terminal of the sail's point it is proceed as follows:

a. The card sends any message to the center .

b. The center sends a random message E back to the card .

c. The card computes $\gamma_S(E)$ and gives the result back .

d. The center computes $\gamma_S(E)$ by means of its card and compares it with the customer's result .

e. The communication is refused if the results are both different .

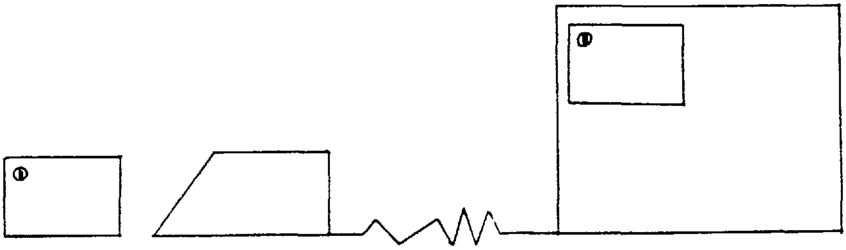


figure 1.

This is just a simple example; but it is possible to design a similar protocol where each card has a different secret key, i.e. depending on a P.I.N., N . Here, the card has to send N in the phase a of the protocol . It is also possible to add a password for the user , so that one can authenticate a valid card and its legal user .

Let's remark that all these protocols don't use the inverse function of γ, if it exists .

2. Introduction to the functions studied in this paper .

We have just seen that it is important to device one-way encryption functions for smart cards . The problem with these new objects is that their ROM, wich contains the software (i.e. the code of the encryption function γ), is relatively small . We have to design "simple" functions that are "complicated" enough to be secure ! On the other side, the RAM is exceedingly small. We are interested in algorithms that, at any step of the computations, load the smallest part of the data as possible in this RAM .

Here, we describe a family of functions that satisfy these requirements. We shall see that these functions can accept very big keys. Therefore we shall study the security of these functions in term of the length of these keys .

II. THE FUNCTION

1. Description .

The set $\mathbb{M}$ of messages that we consider includes the plaintext and enciphered messages . Let k be a small number (i.e. k = 8) and let $\mathbb{K}$ be the set of k-bits blocks .

We first start considering a family of functions ω_K, $K \in \mathbb{K}$, from $\mathbb{M}$ to $\mathbb{M}$. These functions are supposed to be easily implemented and they use a little key, but we do not ensure that they are bijective.

Now we choose an integer n , and $S = (K_1,\dots,K_n) \in \mathbb{K}^n$ and we set

$$\gamma_S = \omega_{K_n} \circ \dots \circ \omega_{K_1} \quad .$$

The implementation of γ uses a little more instructions that the ω's one and its key can be as large as desired . Moreover, each step of the computation of $\gamma(E)$, $E \in \mathbb{M}$, consists of a calculation of ω and so, uses only a small part of the key S .

2. The problem .

We have to keep in mind that the user of a card will eventually compute as many encrypted messages $\gamma_S(E)$, $E \in \mathbb{M}$, as he wishes . So a first question is :

"Is it possible to guess the secret key S from a big amount of couples $(E, \gamma_S(E))$ where $E \in \mathbb{M}$?" .

However, it is easy to request that the cards compute $\gamma_S(E)$ if and only if E has a definite standard form (i.e. the last byte of E represents the current year modulo 2^8). Then, the opponent user will have to choose the messages E in a subset $\mathbb{M}'$ of standard messages of $\mathbb{M}$. In fact, it will be seen that, if the functions ω are not bijective and satisfy some reasonable hypotheses, the following question has an affirmative answer :

"Is it possible to guess the secret key S by means of a big amount of encrypted messages $\gamma_S(E)$ where $E \in \mathbb{M}'$?" .

3. Notations and hypotheses .

The notations concerning $\mathbb{M}$, $\mathbb{M}'$, $\mathbb{K}, \omega_K$ and γ_S are kept on .

As ω is a "little" cipher function, its program uses a "small" flow chart. It is therefore possible to go backward on it so that we have a relatively fast algorithm, with a mean running time of T, wich gives for each $(F,K) \in \mathbb{M} \times \mathbb{K}$ the list of all the E in $\mathbb{M}$ such that $\omega_K(E) = F$. By means of a spanning tree algorithm, we determine and define the following numbers and sets :

$$\begin{aligned} A_K(F) &= \{E \in \mathbb{M}, \omega_K(E) = F\} , \\ A_{K_m,\dots,K_1}(F) &= \{E \in \mathbb{M}, \omega_{K_m} \circ \dots \circ \omega_{K_1}(E) = F\} , \\ X_K(E) &= \text{Card } A_K(E) , \\ X_{K_m,\dots,K_1}(E) &= \text{Card } A_{K_m,\dots,K_1}(E) . \end{aligned}$$

The functions X_K and $X_{K_m,\dots,K_1}$ have to be considered as random variables on the set $\mathbb{M}$. The hypotheses that are given below give a reasonable model for γ .

(H1) The random variables X_K, $K \in \mathbb{K}$, are pairwise independent and have the same law as an integer random variable Y .
We set, for every $n \in \mathbb{N}$, $p_n = \text{Prob } (Y = n)$.

(H2) The number p_0 is not null .

(H3) Let K, L be in $\mathbb{K}$. Let's suppose that F and G are randomly and independently chosen in $\mathbb{M}$ so that $\omega_K(F) = \omega_K(G) = E$. On the one hand, there are no a priori correlations between $X_L(F)$ and $X_L(G)$; on the other hand there is no correlation between $X_L(F)$ and $X_K(E)$.

(H4) Let $K_1,\dots,K_m,L$ be elements of $\mathbb{K}$. If E is chosen randomly in $\mathbb{M}$ and if F is selected at random in $A_{K_m,\dots,K_1}(E)$, there is no correlation between $X_L(F)$ and $X_{K_m,\dots,K_1}(E)$.

(H5) It is recalled here that the smart card accepts to perform the computation of $\gamma_S(E)$ if and only if $E \in \mathbb{M}'$. Moreover it is assumed that the law of X_K, $K \in \mathbb{K}$, is the same on $\mathbb{M}'$ as on $\mathbb{M}$.

(H5) This is a technical hypothese that says that p_1 is not null .

Some remarks .

a. The hypotheses H2, which is obviously satisfied when ω is not bijective, gives the start point of the search of the key S . Let F , in $\gamma_S(\mathbb{M})$,be such that $X_L(F) = 0$, where $L \in \mathbb{K}$; then one clearly has $L \neq K_n$. That gives a way to through little blocks of keys away .

b. The hypotheses H1, H3, H4 are sound because ω is presumed to mix up the bits contained in E and K . Moreover, an unsuspected dependence , which could invalidate one of these hypotheses, could as well result in a new way of attack against the secrecy of S .

c. The hypothese H6 is not fundamental but the search of S is faster with it .

d. The main task of H5 is to prevent the kind of attack where one chooses many E in $\mathbb{M}$ and then analyses the so obtained couples $(E, \gamma_S(E))$. It is possible to show that, if H5 is false, such a search of S is very easy for the owner of a card .

4. Study of the random variables $X_{K_m,\dots,K_1}$.

It is deduced from the hypotheses H1, H3, H4 that for every $m \in N^*$ and $(K_1,\dots,K_m) \in \mathbb{K}^m$, the random variables $X_{K_m,\dots,K_1}$ have the same law as a variable that will be noted Y_m, from now on . The law of Y_m is deduced from Y_{m-1} as follows :

"Y_m is the sum of Y_{m-1} independent variables with the same law as Y_1" .

Let f be the generating function of Y_1 and f_m be the one of Y_m . The end of the paragraph III.15 of "Calcul des probabilités" of Renyi (Dunod) gives :

Proposition 1. The functions f_m satisfy $f_{m+1} = f_m \circ f$, for every $m \in \mathbb{N}^*$.

Let σ be the standard variation of Y and, for every r.v. Z, let $E(Z)$ and $\sigma(Z)$ stand respectively for the mean value and the standard variation of Z . The proposition 1 allows us to study Y_m and we obtain the following assertions :

Proposition 2.

a. For every $m \in \mathbb{N}^*$, one has $E(Y_m) = 1$ and $\sigma(Y_m) = \sqrt{m}\,\sigma$.

b. The series Prob $(Y_m = 0)$ has 1 as limit when $m \to \infty$ and

$$1 - \text{Prob}\ (Y_m = 0) \sim 2/\sigma^2 m \quad .$$

c. For every $\varepsilon > 0$, there is a constant $C > 0$ such that, for every $m \in \mathbb{N}^*$ one has :

$$\text{Prob}\ (Y_m = 1) > C/m^{2+\varepsilon} \quad .$$

Demonstration .

First we compute $E(Y_1)$. To do this, one has to compute $E(X_K)$ where K is any element of $\mathbb{K}$. The numbering of $\mathbb{M}$ gives

$$\text{Card } \mathbb{M} = \sum_k k.(\text{Prob}(X_K = k)\ .\ \text{Card } \mathbb{M}) \quad ,$$

and $\sum k.p_k = 1$. Therefore we have $E(Y_1) = 1$. Then, it is easy to see that

$$E(Y_m) = E(Y_1)^m = 1 \quad .$$

The calculation of $\sigma(Y_m)$ is done recursively on m . We use the following formulas :

$$\sigma^2(Y_1) = f''(1) + f'(1) - f'(1)^2 = f''(1) \quad ,$$

$$f'_m(1) = 1 \quad ,$$

$$f''_m(z) = f''(z).f'_{m-1}(f(z)) + f'(z).f''_{m-1}(f(z)) \quad ,$$

we obtain :

$$f''_m(1) = \sigma^2 + f''_{m-1}(1) \ ,$$

and one has the part a of the proposition .

Now the series $(f_m(0))$ is strictly increasing and has 1 as limit . In order to obtain b, one shows the following fact:

For every $\varepsilon>0$, there exists an $M \in \mathbb{N}$ such that, for every $m > M$, one has

$$\frac{\sigma^2}{2+\varepsilon}\, m < \frac{1}{1-f_m(0)} < \frac{\sigma^2}{2-\varepsilon}\, m \quad .$$

Let m be an integer and set $a = 1 - f_m(0)$. From the Taylor formula, there is a real $\theta \in]0,1[$ such that

$$f(1 - a) = 1 - a + \frac{a^2}{2} f''(1 - \theta a) \ .$$

Therefore :

$$\frac{1}{1-f_{m+1}(0)} - \frac{1}{1-f_m(0)} = \frac{1}{2} \frac{f''(1-\theta a)}{1-\frac{a}{2} f''(1-\theta a)} .$$

The right hand side is a continuous function of (θ, a) that takes the value $\sigma^2/2$ at $(0, 0)$; therefore there exists an integer m_0 such that the left hand side lies in the interval $]\sigma^2/(2+\varepsilon/2) , \sigma^2/(2-\varepsilon/2)[$ for every $m > m_0$. That gives b to us .

The demonstration of c starts from the following formula ;

$$f_m'(0) = f'(f_{m-1}(0))\dots f'(f_1(0)).f'(0) ,$$

which is a consequence of the proposition 1 . Because of the hypothese H6, this term is not null . One uses the estimation of $f_k(0)$ just given to apply the Taylor formula at $f'(f_k(0))$. One has then to find a lower bound for the following product

$$\prod_{k=b}^{n} (1 - a/k)$$

where $a = 2+\varepsilon$ and $b > 0$. This last product is equivalent to

$$\frac{\Gamma(b+1)}{\Gamma(b-a+1)} n^a$$

(use the Stirling formula) , and one can conclude .

We end this section with the study of the random variables $X_{K_n,\dots,K_{n-m+1}} \circ \gamma_S$, where $m \leqslant n$. Let $Y_m \circ \gamma$ be that variable . It is met when one takes $E \in \mathbb{M}'$ at random and then computes $F = \gamma_S(E)$ by means of the card to observe $A_{K_n,\dots,K_{n-m+1}}(F)$.

Proposition 3. For every r $\mathbb{N}$, one has

$$\text{Prob } (Y_m \circ \gamma = r) = r.\text{Prob } (Y_m = r) .$$

The generating function of the r.v. $Y_m \circ \gamma$ is $z \longmapsto z.f_m'(z)$.

The second assertion is directly deduced from the first one and gives with the proposition 2 :

Corollary .

a. For every $\varepsilon > 0$, there is a constant $c > 0$ such that, for every m , one has

$$\text{Prob } (Y_m \circ \gamma = 1) > c/ m^{2+\varepsilon} .$$

b. $$E(Y_m \circ \gamma) = 1 + m\sigma^2 .$$

In order to show the proposition 3, we have to give an estimation of Prob $(Y_m \circ \gamma = r)$. First we start numbering the elements $x \in \mathbb{M}$ such that $Y_m(\gamma(x)) = r$, which is the same as to sum the $Y_n(y)$ where $y \in \mathbb{M}$ and $Y_m(y) = r$. Therefore

$$\begin{aligned} \text{Prob } (Y_m o \ = r) &= \frac{1}{\text{Card } \mathbb{M}} \sum_{1 \in \mathbb{N}} 1 \sum_{y \ \mathbb{M}, Y_m(y)=r, Y_n(y)=1} 1 \\ &= \sum_{1 \in \mathbb{N}} 1.\text{Prob } (Y_m = r \text{ and } Y_n = 1) \\ &= \sum_{1} 1.\text{Prob}(Y_n = 1 \ Y_m = r).\text{Prob } (Y_m = r) \ . \end{aligned}$$

Let's keep in mind that Y_n and Y_m represent respectively $X_{K_n,\ldots,K_1}$ and $X_{K_n,\ldots,K_{n-m+1}}$. Therefore Y_n is the sum of r independent r.v. with the same law as Y_{n-m} when it is known that $Y_m = r$. The conditional mean value if Y_n is then equal to r when $Y_m = r$. Therefore , we obtain the proposition 3 .

5. The search of the key .

It is shown here how some elements of $\omega_L(\mathbb{M}')$ can be used to find L ; wich directly applies to search K_n . Let $(F_i)_i$ be a random series of elements of $\mathbb{M}$ and, for every $L \in \mathbb{K}$, let J(L) be the smallest index i such that $X_L(F_i) = 0$. The mean value of J(L) is $1/p_0$.

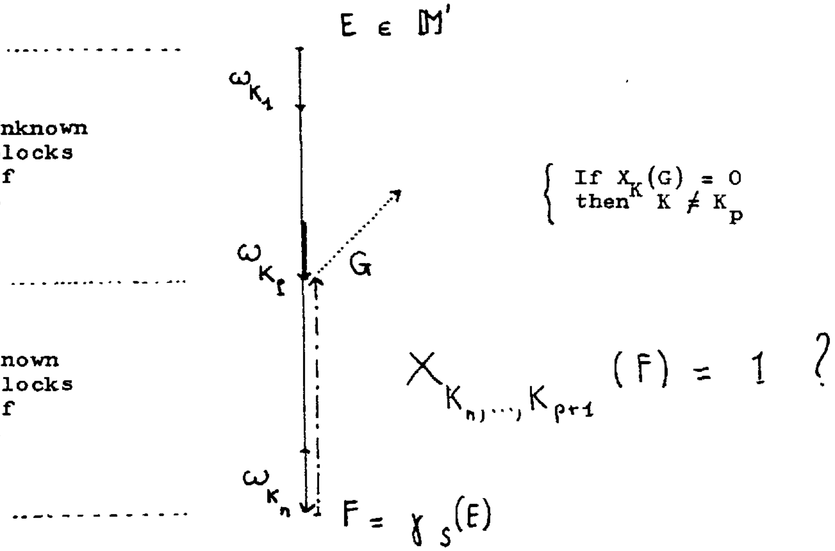
Let h be an integer and $L_1,\ldots,L_h$ be elements of $\mathbb{K}$. We suppose that we have chosen the F_j in $\omega_{L_0}(\mathbb{M}')$, for some unknown $L_0 \in \mathbb{K}$, where $L_0 \neq L_i$ $(1 \leq i \leq h)$. We want to guess that $L_0 \neq L_i$, using the minimum number of F_j . Let J be the maximum of the values $J(L_1),\ldots,J(L_h)$. We know now that, for every L_i, if we compute $X_{L_i}(F_1),\ldots,X_{L_i}(F_k),\ldots$, there will be some j such that $X_{L_i}(F_j) = 0$, with $j \leq J$. So we can eliminate L_i with less than J tries (cf. the remark after the hypothese H2) . If we compute, for every F_i, the vector $(X_K(F_i))_{K \in \mathbb{K}}$, (remember that $\mathbb{K}$ is a relatively small set), J vectors are enough for the elimination of all the L_i $(1 \leq i \leq h)$. An easy computation shows that the mean value of J is O(Log h) .

Now, the search of the key S is done step by step .

To find K_n, as we have $\gamma_S(\mathbb{M}') \subset \omega_{K_n}(\mathbb{M})$, we pick at random messages $F \in \gamma_S(\mathbb{M}')$ and, for every such F, we compute the vector $(X_K(F))_K$. We through each K such that $X_K(F) = 0$ away . This elimination process is complete after O(Log Card $\mathbb{K}$) = O (k) vector computations .

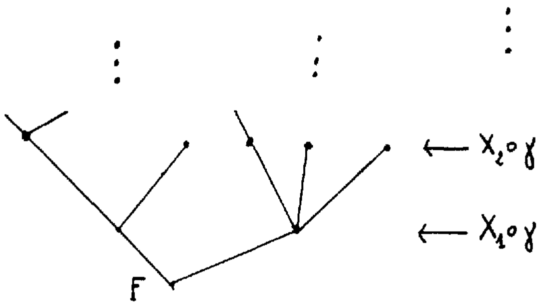
Now let $m \neq n$; the main difficulty here is to find random messages $F \in \omega_{K_{n-m}}(\mathbb{M})$. We proceed recursively and we suppose that $K_n,..,K_{n-m+1}$ are known . First of all, we look for messages $E \in \mathbb{M}'$ such that the set $A_{K_n,\ldots,K_{n-m+1}}(\gamma_S(E))$ has exactly one element F . The corollary of the proposition 3 says that we must perform a mean value of $O(m^{2+\varepsilon})$ tries to find such a message . We then apply the above mentioned method at about O(k) messages F to find K_{n-m} .

Figure 2.



Before giving the result of this paragraph, we have to perform the estimation of the mean number of operations needed to compute $Y_m \circ \gamma_S$ at any $E \in \mathbb{M}$. Let G be the mean number of operations needed to compute ω ; $\gamma_S(E)$ needs about nG operations when computed at any (E, S) in $\mathbb{M} \times \mathbb{K}^n$. Let's remember that the mean number of operations needed for the calculation of Y_1 is T . To obtain $X_{K_n,\dots,K_{n-m+1}}(F)$, one has to develope a tree whose depth is m . The root, of depth 0, is F; the sons of every node of depth p are its antecedents by $\omega_{K_{n-p}}$. Every node F' of depth p induces the calculation of an $A_K(F)$, which means about T operations . The mean number of these nodes is $\sum_{p=1}^{m-1} E(X_p \circ \gamma) = O(m^2)$. Therefore $O(m^2T + nG)$ operations are approximately needed to compute $Y_m \circ \gamma$.

Figure 3.


$\longleftarrow X_2 \circ \gamma$
$\longleftarrow X_1 \circ \gamma$
F

Now, we apply these results and the corollary of the proposition 3 with $\varepsilon = 0.1$ to conclude this paper with the following proposition :

Proposition 4. With the above mentionned hypotheses, there exists an algorithm that allows the owner of a smart card using γ, who is allowed to compute enciphered messages of $\mathbb{M}'$ at will, to find the secret key of γ with a mean number of

$$O\ (\ kTn^{5.1} + Gn^{4.1} + k2^{k}Tn\ )$$

operations .