NON LINEAR NON COMMUTATIVE FUNCTIONS FOR DATA INTEGRITY

S. HARARI
Université de Toulon et du Var
Château Saint Michel
83130 - LA GARDE

1- INTRODUCTION :

Several authors have recently proposed digital signature schemes [1], [2],... In an environment where identification is not possible, and the transmission safe the use of these schemes certify that the data originated from the legitimate person. However in an environment where identification can be ensured by other means and where transmission is done in an unsafe medium, the use of these same schemes ensure data integrity : any modification of the data during transmission shows up when one checks the corresponding signature.

The systematic use of signature functions for data integrity has two important shortcomings :
1) The redundancy introduced by the signature schemes is about as long as the data to be protected.
2) The average number of computing steps per protected digit is very large.

In this paper we introduce some functions allowing the use of data integrity witnesses which introduces minimal redundancy (50 digits for about 10.000 digits of data). The average number of operations per protected digit is kept small.
We study the cryptographic strength of these functions and show that it increases with the length of the data being protected.

II - SEAL FUNCTIONS

Let $\mathcal{T}$ be the set of texts made of strings of h decimal digits : $\mathcal{T} = R^h$ where R is the ring of decimal digits. Let $\mathcal{J}$ be the set of strings of p

T. Beth, N. Cot, and I. Ingemarsson (Eds.): Advances in Cryptology - EUROCRYPT '84, LNCS 209, pp. 25-32, 1985.

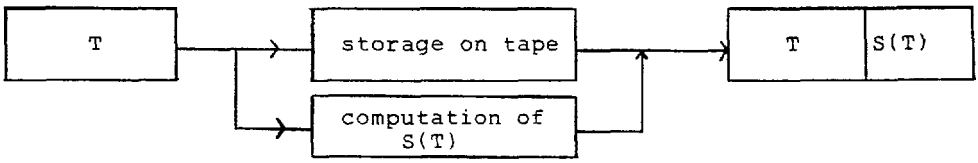
decimal digits : $\mathcal{Y} = R^p$. A seal function s is a function :

$$s : \mathcal{X} \longrightarrow \mathcal{Y}$$

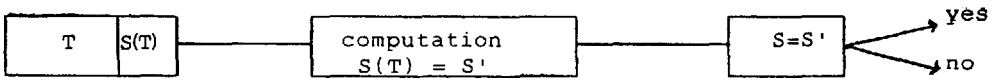
A seal of a text T is then s(T)

The seal function is used or storing in the following way :
Prior to the transmission of the data on an unsafe medium, a seal is computed. It is then processed with the data.

When retrieving the data from an unsafe medium; a seal is recomputed from the data and compared with the one that is retrieved from the unsafe medium. If the two seals coïncide the data is considered free from alterations.



seal function



seal check

III - CONSTRAINTS ON THE SEAL FUNCTION

a) length of the seal

The seal being a decimal quantity, a length p of more then 20 decimal digits is enough to ensure that random attacks on the seal have a low probability of succeeding. If n is any integer between 10.000 and 100.000, then one is sure that in any application the data flow is not interrupted too often for seal computation or recomputation.

b) attacks on the seal and unforgeability

The data that is to be protected by the seal is highly structured. The structure and content is known to an opponent. The aim of an opponent

to the system is to modify the data, and if necessary (and possible) the seal, in such a way that the modified seal is legitimately associated to the modified data.

Let n = 10.000 and p = 20

Let $s = R^{10.000} \longrightarrow R^{20}$ be a seal function.
For a given seal S, the cardinality of S^{-1} (S), that is the set of texts having a given seal is of the order of R^{9980}.
Any structure in that set, will help an opponent in finding many of its elements.
Any structure relating $S^{-1}(S)$ and $S^{-1}(S')$ for two different seal S and S' will also help an opponent in finding many of its elements.
This leads us to the following conditions.

i) The mapping

$s : R^n \longrightarrow R^p$ is a random variable, uniformely distributed on R^p for each probability distribution on R^n.

ii) For any given text $(t_1,\ldots,t_n)$. Let $I = \{1,\ldots,n\}$
The mapping :

$$I \times R \longrightarrow R^p$$

$$(i,r) \longrightarrow s\ (t_1,\ldots,t_{i-1},\ t_i+r,\ t_{i+1},\ldots,t_r) - s\ (t_1,\ldots,t_n)$$

Should be a uniformely distributed random variable, for each probability distribution on I x R.

iii) Let S_n be the permutation group of $I = \{1,\ldots,n\}$, and δ an element of S_n, for any given text $(t_1,\ldots, t_n)$ the mapping

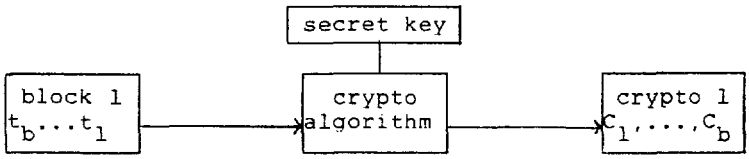
$$S_n \longrightarrow R^p$$

$$\delta \longrightarrow s\ (t_{\delta(1)},\ldots,t_{\delta(n)}) - s\ (t_1,\ldots,t_n)$$

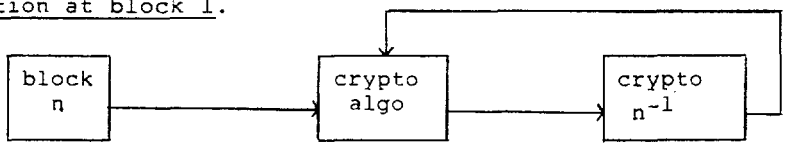
should be a uniformely distributed random variable for each probability distribution on S_n.

c) computational complexity of seal function
A seal function is primarly intended to be used in software. Therefore

a seal function should have a low computational complexity per protected digit. Most of the well known cryptographic algorithms have a high computational complexity ciphered digit of data and therefore perform poorly in software. Using a cryptographic algorithm in a feedback mode, meets the unforgeability requirements, but leads to a very slow seal computation.



First computation
computation at block 1.



computation at block n.

IV - SOME EXAMPLES OF SEAL FUNCTIONS

a) the sum function

Let b be any integer between 1 and n :

$$\text{For a text } (t_1, \ldots, t_n)$$

Let $T_1 = t_1 t_2 \ldots t_b$, $T_2 = t_{b+1} \ldots t_{2b}, \ldots, t_n = t_{(n-1)b+1} \ldots t_{nb}$

Define the seal of the text $t_1 \ldots t_n$ as

$$S(t_1, \ldots t_n) = \sum_i T_i$$

This seal depends on every digit of the text, but does not satisfy neither requirement i) nor ii). Any permutation of the digits of the text corresponding to the permutation of blocks on the T_i's lead to the same seal.

the same seal.

b) the sum of cryptos

Let C be a cryptographic function, assigning to each set of b integers of text, a cryptogram of length d

$$C : \mathbb{Z} \longrightarrow \mathbb{Z}$$
$$C : R^b \longrightarrow R^d$$

$$(t_1,\ldots,t_b) \longrightarrow (c_1,\ldots,c_d) = c\,(t_1,\ldots,t_b)$$

let $C_1 = c_1 \ldots c_d = c(t_1,\ldots,t_b)$

$C_2 = c_{d+1}\ldots c_{2d} = c(t_{b+1},\ldots,t_{2b})$

$\ldots$

$C_n = c_{(n-1)d+1}, c_{nd} = c(t_{(n-1)b+1}\ t_{nb})$

$\ldots$

and let $S(t_1,\ldots,t_n) = \sum_i c_i$

This seal function satisfies requirement i) but not ii). Any permutation of the digits of the text corresponding to the permutation of blocks being ciphered lead to the same seal.

c) the concatenation of signatures

Let r,q be two large primes, kept secret and m = q.r
Let l be the length of m; and $T_1 = t_1 t_2 \ldots t_l$, $T_2 = t_{l+1}\ldots t_{2l}$, ..., $T_n = t_{(n-1)l+1},\ldots t_{nl},\ldots$

The legitimate owner of the text knowing the factorisation of m, can easily compute square roots in $\mathbb{Z}_m$.

Let $s_1 = \sqrt{T_1} \bmod m$, $s_2 = \sqrt{T_2} \bmod m, \ldots, s_n = \sqrt{T_n} \bmod m$,

Define $s(t_1,\ldots,t_n)$ as $(s_1,\ldots,s_n)$

This seal function meets most of the requirements on unforgeability. Its shortcoming is the computational effort as well as the length of the seal which is as long as the text itself; and a permutation of two portions of the text as well as their corresponding signature, lead to a new legitimate altered seal.

d) a new seal function (1)

Let A = (aij) be square matrix in order n, whose entries are random integers, chosen by the originator of the text and kept secret.

Let $T = (t_1,\ldots,t_n)$ be a text

$$s(T) = T^t A T = \sum_{i<j} a_{ij} \cdot t_i \cdot t_j$$

is the seal of T

This seal involves only arithmetic operations for its computation. The total number of operations to compute a seal is seen to be $O(n^2)$ multiplications and $O(n^2)$ additions. It is easily checked that the unforgeability requirements are met.

A potential forger of seals has to know the matrix a_{ij} in order to create a legitimate seal for a given text.

Let us suppose that the forger holds

u texts $(t_1^{(1)},\ldots,t_n^{(1)}),\ldots,(t_1^{(u)},\ldots,t_n^{(u)})$ and their corresponding

seals $s_1, s_2,\ldots,s_u$.

To obtain the matrix (a_{ij}) he has to solve the following system

$$\sum_{i,j} a_{ij}\, t_i^{(1)} t_j^{(1)} = s_1$$

$$\sum_{i,j} a_{ij}\, t_i^{(u)} t_j^{(u)} = s_u$$

In this system the t_i, t_j, and $s^{(u)}$ are knowm.

Two methods are indicated to make this system unsolvable.

- Change the matrix (a_{ij}) often enough so that a potential forger cannot obtain enough information from existing seals in order to solve the system.

- Choose n large enough, so that the best known algorithms for solving dense linear systems fail to do so in a short amount of time. The seal system is strengthened by increasing the length of the text being protected.

e) a new seal function (2)

Let $A = (a_1,\ldots,a_{n'})$ be a sequence of random integers of length n', chosen the originator of the text and kept secret

Define $$S(T) = A^t\ T\ A = \sum_{i<j} a_i\ a_j\ t_{ij}$$

This seal function involves only arithmetic operations to compute a seal, and the total number of operations is then $2n'^2$ additions and $2n'^2$ multiplications. It meets the unforgeability requirements.

A potentail forger has to know the sequence A in order to create a legitimate seal for a given text.

Let us suppose he knows u texts $(t_1^{(1)},\ldots,t_n^{(1)}),\ldots,(t_1^{(u)},\ldots,t_n^{(u)})$ and their corresponding seals $s_1,\ldots,s_u$.

He therefore has to solve a system, which is quadratic in the unknow a_i's.

$$\sum_{i,\ j} a_i\ a_j\ t_j^{(1)} = s_1$$

$$\sum_{i,\ j} a_i\ a_j\ t_j^{(u)} = s_u$$

The complexity of finding the a_i's is equivalent to factoring this polinominal.
The complexity of this problem is O (n^3 + log n n^2) in a modular version of the problem.

ooo OOOO ooo OOO ooo

REFERENCES :

[1] Ong schnorr Shamir. An efficient signature schemes based on quadratic equations. 16th Symposium on the theory of comp. 1984, Washington.

[2] Rabin M.O. Probabilistic algorithms in finite fields. Siam J on Computing 9 (1980) 273-280.

[3] S. HARARI. Functions in transmission and storage.
NATO-ASI. The impact of Processing techniques on communications.
Château de Bonas 11-12 july 1983.

[4] Knuth - The art of computer programming seminumerical algorithms
Vol.2 Sec. Ed. 1980, Addison Wesley.