2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Non-Linearity and Security of Self Synchronizing Stream Ciphers

Philippe Guillot† and Sihem Mesnager‡

†‡ Université Paris 8, Vincennes-Saint-Denis,
2, rue de la Liberté, 93 526 Saint-Denis Cedex,
Email: †philippe.guillot@univ-paris8.fr, ‡hachai@math.jussieu.fr

Abstract—Several proposed chaos based ciphers exploit the ergodic property of chaotic orbits. As chaotic systems are unstable and have sensitive dependence on initial conditions, the main difficulty for the receiver is to reproduce the chaotic signal that has been generated by the sender in order to correctly decrypt the message. This is performed by a self synchronizing device. In discrete cryptography, the closest scheme is the so called self synchronizing stream cipher (SSSC). After recalling general security models for assessing cryptographic algorithms, we present SSSC scheme and two examples of cryptanalysis. In order to resist to theses attacks, the ciphering function must satisfy high non-linearity properties which are presented.

1. Introduction

The goal of cryptography is to insure confidentiality and authenticity of information. This is performed by a public ciphering function that involves a secret key in a certain mode of operation. The union of the ciphering function and the mode of operation constitutes the *cryptographic system*.

The three actors of a cryptographic system are the sender, the receiver and the adversary. In a symmetric ciphering system, the plain text is encrypted into a cryptogram using a secret key shared by the sender and the receiver. The goal of the adversary is to get information about the plain text. This can be done by recovering the secret key, but there may exist other ways, in particular to retrieve partial information.

A cryptographic system is called *unconditionally secure* if the adversary has no better strategy than choosing the plain text at random. In the practical world, the security level is assessed taking into account the protection which is supposed to be insured, the data and computation power at the disposal of the adversary. Knowledge of only the cryptogram is not sufficient to decrypt.

In the so called *chosen plain text attack*, the adversary is supposed to know the cryptograms that correspond to messages of his choice.

A cryptographic system is called *semantically secure* if the adversary cannot distinguish the cryptogram from a pure random sequence with reasonable amount of time and computation power.

Finally, it is admitted by the cryptographic community that a ciphering function is secure if the adversary has no better strategy than trying all the possible secret keys in a chosen plain text attack.

Cryptographic algorithms design is still based on confusion and diffusion principles stated by Shannon in 1949 (see [14]).

Diffusion means that a bit change in the key is propagated in the whole ciphertext. It is generally performed by linear transformations.

Confusion means that the relationship between the key, the plaintext and the ciphertext is complex and involved. It is performed by nonlinear transformations, implemented as Boolean functions.

These principles are very close to randomness and the sensitivity to initial conditions that characterize chaos. This makes natural the idea of using of complex dynamic systems for cryptographic applications.

2. Self-Synchronizing Stream Cipher

In a chaos based cryptographic system, the information is hidden by addition of a chaotic signal. This presents strong similarities with conventional stream ciphers. The same chaotic signal has to be reproduced by the receiver despite high sensitivity to initial condition. This is performed by re-synchronizing the chaotic generator by the received signal itself. This is very similar to what is done in *self-synchronizing* stream cipher presented in this section.

In a conventional binary stream cipher, each plain text symbol m_t is combined with a key stream symbol k_t by a group operation to define the cipher text symbol c_t:

$$\forall t \in \mathbb{N} \quad c_t = m_t \oplus k_t \tag{1}$$

For deciphering, the inverse operation is performed with the same key stream symbol. Most of the times, symbols are binary. The group operation is the *exclusive or*, and the inverse operation is the same.

The various stream ciphers are classified depending on how the key stream is generated.

This principle is very old. It was proposed by Vigenère (1586) for the Latin alphabet with a cyclic key stream. As it was first cryptanalyzed by Babbage in 1854, this is for the moment the longest resistant cryptographic system.

In the Vernam stream cipher (1919), the key stream is a pure random sequence. This is the only known system for which the unconditional security is proved. The sender and

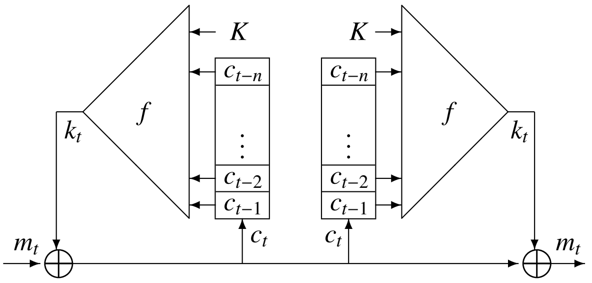

Figure 1: Self Synchronizing Stream Cipher model

the receiver have to share an amount of data as large as the message to encrypt.

The synchronous stream cipher simulates the Vernam system with a pseudo-random key stream generated by both the sender and the receiver, from a shorter seed that constitutes the secret key K. The key stream symbol k_t only depends on the time t and on the secret key K. In order to insure basic security feature, the key stream must be statistically indistinguishable from a true random sequence and must be unpredictable.

In a Self-Synchronous Stream Cipher (SSSC), the key stream only depends on the key and on a bounded number of the last cipher text symbols. It may be represented by a keyed Boolean function $f : \mathbb{F}_2^n \times \mathbb{F}_2^k \to \mathbb{F}_2$, called the *ciphering function*, of n bits of the cipher text $(c_t)_{i\in\mathbb{N}}$ and of the key K as shown on figure 1.

$$\forall t \in \mathbb{N} \quad k_t = f(c_{t-1}, \ldots, c_{t-n}, K) \qquad (2)$$

If the n last cipher text symbols are properly received, then the key stream symbol is correctly generated. This allows the receiver to lose data and to automatically resynchronize the key stream after correctly receiving n symbols.

The scheme presented in figure 1 is a model. The ciphering function f is needed to be complex and is generally implemented as a composition of several rounds of simpler functions that each involves a sub-key build from the secret key K.

This SSSC is the closest conventional scheme from those proposed with a chaotic key stream generation (see for example [7]). The conventional SSSC cryptanalysis may constitute a starting point to evaluate the security of chaos based schemes.

3. Attacks Against SSSC

The core of an SSSC is the keyed ciphering function f. It can be proved that the SSSC is secure as long as this keyed function behaves like a random Boolean function while the key is a random variable. This implies that the key is deeply involved in the definition of f.

A known plain text corresponds to the knowledge of couples $(x, f(x, K))$, where x is a n-dimensional vector extracted from a length n window in the cipher text. Indeed, let $x_t = (c_{t-1}, \ldots, c_{t-n})$, the value $y_t = f(x_t, K)$ is deduced from relations (1) and (2) by $y_t = m_t \oplus c_t$.

The value of n must be large enough that any practicable known plain text leads to the knowledge of a negligible part of the truth table of f.

The attacks presented below exploit general weakness of the ciphering function. For attacks that make use of the particular architecture of the ciphering function, see for example [2], [3] and [4].

3.1. Ciphering function reconstruction

This attack does not recover the key K, but gives information on the plain text. The adversary knows a certain number of couple $(x, f(x, K))$. If the ciphering function $\varphi : x \mapsto f(x, K)$ has a low algebraic degree, then it can be entirely recovered which allows complete decryption.

This attack is modelled as a decoding problem on a virtual erasure channel.

Let us recall that any n-variable Boolean function is uniquely represented as a n-variable polynomial where, as $x_i^2 = x_i$ for Boolean values, the degree of each variable is at most 1. The algebraic degree of the Boolean function is the degree of this multivariate polynomial. The Reed-Muller code $\mathcal{RM}(n, r)$ is the linear code of length 2^n and dimension $1 + \binom{n}{2} + \cdots + \binom{n}{r}$ constituted of n-variable Boolean functions of algebraic degree less than or equal to r.

The virtual channel consists in sending the truth table of the ciphering function, namely an element of $\mathcal{RM}(n, r)$. But the adversary only receives values that correspond to the known plain text. The others values are erasures in this virtual transmission.

If the adversary can decode, *i.e.* if he can retrieve on demand the missing values of the function φ, then he can decrypt any text ciphered with the key K.

If the ciphering function φ is of algebraic degree higher than r, then the previous attack is still possible but it will only build an approximation of φ that will lead to an estimation of the plain text.

A bound on the remaining errors while decrypting by such a way is given by the so called *covering radius* of Reed-Muller code, which is by definition the larger distance of any Boolean function from the code (see [1]).

3.2. General linear attack

The linear attack is a known plaintext attack that has been first published by Matsui in 1991 for cryptanalyzing the DES ([9]). A variant of this attack may be applied to SSSC without any assumption on the design of the ciphering function. This attack recovers the secret key.

A linear approximation of the ciphering function f is a triple $(\alpha,\beta,\varepsilon) \in \{0,1\}^n \times \{0,1\}^k \times \{0,1\}$ such that, for any n-dimensional binary random vector X, the probability

$$P\Big[\alpha \cdot X + \beta \cdot K + \varepsilon f(X, K) = 0\Big]$$

is significantly different from 1/2.

The adversary must know $L > 1$ linear approximations $(\alpha_i, \beta_i, \varepsilon_i)_{i \in \{1,\dots,L\}}$.

The key K is unknown, but several couple $(x, f(x, K))$ are known.

For each values of the known couples $(x, f(x, K))$, the adversary count the number of times $\alpha_i \cdot x + \varepsilon_i f(x, K) = 0$. If this occurs more that one in two times, then he decides that $\beta_i \cdot K = 0$, otherwise he decides that $\beta_i \cdot K = 1$. This gives a linear relation $\beta_i \cdot K = \eta_i$, satisfied by the key K.

The number of known plain text needed depends on the quality of the linear approximation in order to take the right decision.

This process has to be repeated for several other linear approximations. Once a sufficient number of such relations is established, a matrix inversion retrieves the key K.

4. Non-Linearity Criteria for Boolean Functions

The previous section shows that the ciphering function must have good non linearity properties in order to resist the presented attacks (see [10]). An elegant mathematical tool to analyse the cryptographic properties of a Boolean function f is its Walsh transform, which is by definition the Fourier transform of its sign function $\chi_f = (-1)^f$. Let $\mathbb{F}_2^n = \{0, 1\}^n$. The Hamming weight of a vector in $\mathbb{F}_2^n$ is the number of its non-zero coordinates. Let f be a n-variable Boolean function on $\mathbb{F}_2^n$, the Walsh transform of f, denoted by $\widehat{\chi_f}$, is by definition

$$\widehat{\chi_f}(\omega) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+\omega \cdot x},$$

where $\omega \cdot x = \omega_1 x_1 + \cdots + \omega_n x_n$. A *Fast Fourier Transform* algorithm computes $\widehat{\chi_f}$ with complexity $O(n2^n)$, making this tool effective.

4.1. Distance to the set of affine functions

A first measurement of non linearity of a n-variable Boolean function f is the number of values in the truth table which must be changed to reach the closest affine function. This number is called its *non-linearity* of f. It can be expressed by means of the Walsh transform as $2^{n-1} - \frac{1}{2} \max_{\omega \in \mathbb{F}_2^n} \left|\widehat{\chi_f}(\omega)\right|$.

The non-linearity of Boolean functions used in a stream cipher must be high since the existence of affine approximations allows to attack the cryptosystem as shown in section 3.2. For even n, the maximum possible non linearity achievable by a n-variable Boolean function is $2^{n-1} - 2^{\frac{n}{2}-1}$. Boolean functions that reach this bound are called *bent* functions. However bent functions only exists in even dimension and are not balanced, namely the output of a bent function does not behave as a uniform random variable if the inputs do. A Boolean function f is balanced if and only if $\widehat{\chi_f}(0) = 0$. The maximum possible non linearity of a balanced boolean function is known only for $n \leq 8$. For the larger values of n, the non linearity lies between $2^{n-1} - 2^{\frac{n-1}{2}}$ (which can be always achieved by functions of algebraic degree 2) and $\left\lfloor\left\lfloor 2^{n-1} - 2^{\frac{n}{2}-1} \right\rfloor\right\rfloor$ (where $\lfloor\lfloor x \rfloor\rfloor$ denotes the greatest even number less than or equal to x). The problem of constructing highly non linear Boolean functions is one of the most challenging problems in symmetric cryptography ([5, 12]).

4.2. Correlation immunity

A n-variable Boolean function f is *correlation immune* of order t if all the sub-functions obtained from f by keeping any t input variables constant, have the same distribution. This concept was introduced by Siegenthaler ([15]) in order to resist some cryptanalytic attacks on stream cipher. A Boolean function f is said to be *t-resilient* if it is balanced and correlation immune of order t. The resiliency is characterized by means of the Walsh transform: a Boolean function f is t-resilient if and only if its Walsh transform vanishes for every vector of Hamming weight less than or equal to t. The order of resiliency of a Boolean function used in a stream cipher should be high. However, non-linearity and order of resiliency can not be optimized simultaneously. Indeed, the non-linearity of t-resilient Boolean functions for $t > \frac{n}{2} - 2$ is upper bounded by $2^{n-1} - 2^{t+1}$ and there exists constructions of Boolean functions achieving this upper bound (e.g. see [11]). However the situation is not so clear for $t \leq \frac{n}{2} - 2$. Recently, 1-resilient Boolean function with very high non-linearity $2^{n-1} - 2^{\frac{n}{2}} + 2^{\frac{n}{2}-2} - 2$ has been constructed (see [8]).

4.3. Propagation criterion

A n-variable Boolean function f is said to satisfy the propagation criterion with respect to a vector α if the Boolean function $x \mapsto f(x) \oplus f(x + \alpha)$ is balanced. More generally, a Boolean function satisfies the propagation criterion at order k if it satisfies the propagation criterion with respect to any vector of Hamming weight ranging from 1 to k. The order k of the propagation characteristic of a Boolean function is related to its strength to resist the differential cryptanalysis and so should be high. Walsh transform helps to study the propagation characteristics of a Boolean function thanks to this characterization : a n-variable Boolean function f satisfies the propagation criterion of order k if and only if, for every non-zero vector a of Hamming weight less than or equal to k,

$$\sum_{\omega \in \mathbb{F}_2^n} (-1)^{a \cdot \omega} \left(\widehat{\chi_f}(\omega)\right)^2 = 0.$$

The connection among the various non-linearity criteria is a topic in the area of designing and analysing cryptographic functions. In particular, the quantitative relationship between propagation characteristics and non-linearity, two critical indicators of the cryptographic strength of a Boolean function, have been investigated in [13]. For example, it was pointed out that if a Boolean function f in

n variables satisfies the propagation criterion with respect to all but a subset $\mathfrak{R}$ then the non-linearity of f is greater than or equal to $2^{n-1} - 2^{\frac{n}{2}-1} |\mathfrak{R}|^{\frac{1}{2}}$ (where $|\mathfrak{R}|$ denotes the cardinality of $\mathfrak{R}$). Further improvements have been done and it has been shown that it is possible to construct highly non-linear Boolean functions with very good propagation characteristics.

4.4. Linear structures

A vector α is said to be a linear structure of a Boolean function if the Boolean function $x \mapsto f(x) \oplus f(x + \alpha)$ is constant. The set of linear structures for a given function constitutes a vector-space $\mathcal{L}$. Moreover, assuming that $f(0) = 0$, one has the identity $f(x + \alpha) = f(x) \oplus f(\alpha)$ for every $x \in \mathbb{F}_2^n$ and $\alpha \in \mathcal{L}$. Therefore, on the subspace $\mathcal{L}$, the Boolean function f behaves as a linear function. If $\mathcal{L}$ is non-trivial, then the function f is degenerate, namely it can be replaced by a simpler function of strictly less variable up to an affine isomorphism.

The number of values in the truth table of f which must be changed to reach the closest function that do have a non zero linear structure constitutes another important nonlinearity measure (see [10]).

4.5. High order non-linearity

The non linearity of order r generalises the standard non linearity. It equals the number of values in the truth table which must be changed to reach the closest Boolean function of degree less than or equal to r. This parameter measures the ability to resist to low-degree approximation attacks as those exposed in [6]. These attacks pose a threat even when Matsui's advanced linear cryptanalytic attacks are rendered impractical. High order non linearity of a given Boolean function is difficult to compute. There exists neither an equivalent of the expression of the standard non linearity by mean of the Walsh transform nor a simple algorithm to compute it.

Boolean function with such high order non linearity are difficult to construct. Much less is known about the maximum possible achievable non linearity of order r. The main result is an upper bound (see [1]).

References

[1] C. Carlet and S. Mesnager. "Improving the upper bounds on the covering radii of Reed-Muller codes", to appear in the proceedings of ISIT 2005.

[2] J. Daemen, R. Govaerts and J. Vandewalle, "On the Design of High Speed Self-Synchronizing Stream Ciper", *Singapore ICSS/ISITA '92 Conference Proceedings*, pp 279–283, 1992.

[3] J. Daemen and P. Kitsos, "The Self Synchronizing Stream Cipher MOSQUITO", *Submission to ECRYPT*, 2005.

[4] A. Joux and F. Muller, "Loosening the KNOT", *FSE 2003, LNCS Vol. 2887/2003*, pp 87–99, 2003.

[5] K. Khoo and G. Gong. "New Constructions for Resilient and Highly Nonlinear Boolean Functions", *ACISP*, pp. 498-509, 2003.

[6] L.R. Knudsen and M.J.B. Robshaw. "Non-linear approximations in linear cryptanalysis". *Advances in Cryptology –Eurocrypt'96*, volume 1070 of *LNCS*, pp. 224–236. Springer Verlag, 1996.

[7] L. Larger, J.-P. Goedgebuer, and F. Delorme, "Optical encryption system using hyperchaos generated by an optoelectronic wavelength oscillator", *Phys. Rev. E, vol. 57, no. 6*, pp. 6618–6624, June 1998.

[8] S. Maitra, "On Nonlinearity and Autocorrelation Properties of Correlation Immune Boolean Functions", *J. of Information Science and Engineering, vol. 20*, pp. 305–323, 2004.

[9] M. Matsui, "Linear cryptanalysis method for DES cipher", *EUROCRYPT 93, LNCS vol. 765*, pp. 386–397, 1994.

[10] M. Meier and O. Staffelbach, "Nonlinearity criteria for Cryptographic Functions", *EUROCRYPT 89, LNCS vol. 473*, pp. 549–562, 1990.

[11] E. Pasalic, S. Maitra, T. Johansson and P. Sarkar, "New constructions of Resilient and Correlation Immune Boolean functions Achieving Upper Bound on Nonlinearity", *Workshop on Coding and Cryptography*, Electronic Notes in Discrete Mathematics. Elsevier, January 2001.

[12] P. Sarkar and S. Maitra. "Modifications of Patterson-Wiedemann Functions for Cryptographic Applications", *IEEE Trans. Inform. Theory, vol 48*, pp. 278–284, 2002.

[13] J. Seberry, X. M. Zhang and Y. Zheng. "The relationship between propagation characteristics and nonlinearity of Boolean functions", *J. of Universal Computer Science, vol. 1(2)*, pp. 136–150, 1995.

[14] C.E. Shannon, "Communication theory of Secrecy System", *Bell Sys. Tech journal* vol *28, pp. 656–715*, 1949.

[15] T. Siegenthaler. "Correlation-immunity of nonlinear combining functions for cryptographic applications", *Information Processing Letters, vol. 69*, pp. 76–780, 1984.

[16] R. Tenny and L.S. Tsimring, "Steps Toward Cryptanalysis of Chaotic Active/Passive Decomposition Encryption Schemes using Average Dynamic Estimation", *International Journal of Bifurcation and Chaos, Vol. 14, No 11*, pp. 3949–3968, 2004.