A New Statistical Testing for Symmetric Ciphers and Hash Functions

Eric Filiol*
ESAT - Virology and Cryptology Lab
B.P. 18 35998 Rennes, FRANCE
eric.filiol@esat.terre.defense.gouv.fr

July 23, 2002

Abstract

This paper presents a new, powerful statistical testing of symmetric ciphers and hash functions which allowed us to detect biases in both of these systems where previously known tests failed. We first give a complete characterization of the Algebraic Normal Form (ANF) of random Boolean functions by means of the Möbius transform. Then we built a new testing based on the comparison between the structure of the different Boolean functions Algebraic Normal Forms characterizing symmetric ciphers and hash functions and those of purely random Boolean functions. Detailed testing results on several cryptosystems are presented. As a main result we show that AES, DES Snow and Lili-128 fail all or part of the tests and thus present strong biases.

Keywords: Boolean function, statistical testing, symmetric cipher, randomness, hash function, Möbius transform, Walsh Transform.

1 Introduction

Randomness is the ground property of cryptography. For the attacker, any quantities produced by a given cryptosystem must look as unpredictable as possible. It means that these quantities have to be of sufficient size and "be random" in the sense that the probability of any particular value being selected must be as weak as possible to preclude a cryptanalyst from gaining advantage through optimed search strategy based on such probability [21, p 169].

From a general point of view, any symmetric cipher and any hash function must be designed as a pseudorandom bit generator (PRBG) relatively to each of its output bits.

Two important requirements are then to be satisfied: the output sequences of a PRBG must be statistically indistinguishable from truly random sequences and the output bits must be unpredictable to an attacker with limited computing facilities. Therefore many different statistical tests have been proposed and are usually implemented to evaluate these two requirements. Historically we must cite Golomb's randomness postulates [18]. These tests have been designed as necessary but not sufficient tests to check if a shift register sequence statistically behaves properly.

*also INRIA, CODES Project, Domaine de Voluceau 78153 Le Chesnay Cédex, FRANCE *Eric.Filiol@inria.fr*

Yet statistically good according to these postulates, this kind of sequence has been shown very predictable when using the Berlekamp-Massey algorithm [23]. This is the illustration that randomness is uniquely defined relatively to the statistical tests we may use.

Many other statistical tests have been proposed in order to better improve what may be considered as "random". Among many others, let us cite those that mainly implemented: frequency test, serial test, poker test, runs test and autocorrelation test [2, 14, 20], Maurer's universal statistical test [24], repetition test [17] (for a more detailed bibliography on statistical tests used in cryptography see [21, pp 188-189] and [5]).

To be precise, these tests are primarily intended for stream ciphers (or block ciphers in modes as stream ciphers) whose output sequences are long enough to apply probability results (essentially the central limit theorem). When dealing with pure block ciphers or with hash functions, the scope of these tests could be questionned. In this latter case, the concepts of *diffusion* and *confusion* [30] are generally preferred (it is clear that one could define them from a statistical point of view). However these concepts are defined either rather empirically or too theoretically (for example through equivocation of the key K about the ciphertext Y).

All the recently proposed symmetric cryptosystems and hash functions can be considered as reasonably satisfying all the known randomness requirements (*i.e.* pass all the known statistical tests). Now the essential part of the cryptanalyst's work is to find an exploitable bias, due to an unknown design flaw, that none of the up to now known test detected. For that, the cryptanalyst generally first designs a new hypothesis testing based on a new test. Let us recall that in fact randomness is a theoretical indeed "philosophical" concept. Practically speaking it can only be determined and defined relatively to the set of statistical tests used to evaluate it. Randomness alone is a nonsensical concept.

In this paper we present a new hypothesis testing based on a χ^2 distribution and called Statistical Möbius Analysis. More precisely we define as working statistic the number of monomials of degree exactly d in the *Algebraic Normal Form* (ANF) of all the Boolean functions modeling each of the output bits. The set of these d-monomials effectively present in the ANF are practically computed by means of the Möbius transform. A secure cryptosystem has a fixed distribution determined by general results on random Boolean functions. Then one-sided tests allow us to check if the constituent Boolean functions are truly random.

These tests have been implemented for a few recently proposed stream ciphers and block ciphers, as well as for the main hash functions. All are known to have passed all the previously known statistival tests and thus are considered as having very good random properties. Out main results is that famous cryptosystems such AES, DES and Lili-128 did not pass all or part of our tests.

Section 2 will present the necessary preliminaries and give the characterization of the Algebraic Normal Form (ANF) of random Boolean functions. In particular we complete the results presented in [26], make them more practical and give new results on the total degree of a Boolean function. Section 3 presents the new test we designed whilst Section 4 gives detailed numerical results that have been obtained for a few stream ciphers (Lili-128, Snow, BGML and RC4), block ciphers (DES and AES) and hash functions (SHA-0, SHA-1, Ripe-MD, Ripemd160, Haval, MD4 and MD5). Section 5 finally concludes and presents future work to exploit these biases.

2 Characterization of Boolean Functions and Results

In this section we present a new statistical way of describing Boolean functions by use of their ANF. After the characterization itself with the Möbius transform, we deduce results on the balancedness and correlation properties with the help of the Walsh transform. These results are commented on

in the particular case of Bent functions.

2.1 Structure of the Algebraic Normal Form

A Boolean function is a function f from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. The number of such functions is 2^{2^n}. We define a random Boolean function as a function f whose values are independent, identically distributed random variables that is to say

$$\forall (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n, \quad P[f(x_1, x_2, \ldots, x_n) = 0] = \frac{1}{2}. \tag{1}$$

In other words each $f(x_1, x_2, \ldots, x_n)$ is a *Bernoulli* random variable of parameter $\frac{1}{2}$. The corresponding probabilistic law will be denoted $\mathcal{B}(p)$ whith $p = \frac{1}{2}$ for our present case.

The weight of a Boolean function over $\mathbb{F}_2^n$ is defined by

$$wt(f) = |\{x \in \mathbb{F}_2^n | f(x) = 1\}|.$$

Then a Boolean function will be said to be *balanced* if $wt(f) = 2^{n-1}$. Note that a random Boolean function, as defined above, may be not balanced. In fact we will give the general probability for such a function to be balanced.

The *Algebraic Normal Form* (ANF) of f is the multivariate polynomial given by

$$f(x_1, x_2, \ldots, x_n) = \bigoplus_{u \in \mathbb{F}_2^n} a_u x^u, \qquad a_u \in \mathbb{F}_2$$

where $u = (u_1, u_2, \ldots, u_n)$ and $x^u = \prod_{i=1}^n x_i^{u_i}$. The a_u are given by the Möbius transform of f:

$$a_u = \bigoplus_{x \preceq u} f(x) \tag{2}$$

where $\preceq$ denotes the partial order on the Boolean lattice, that is to say that $\alpha \preceq \beta$ if and only if $\alpha_i \leq \beta_i$ for all $1 \leq i \leq n$. A monomial $a_u x^u$ of the ANF will then be said of degree k if $a_u = 1$ and if $wt(u) = k$ where $wt(.)$ denotes the Hamming weight. With all this notation we now can formulate our first result:

Proposition 1 *The algebraic Normal Form (ANF) of a random Boolean function f from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ has 2^{n-1} monomials in average. For each k such that $0 \leq k \leq n$, there are an average of $\frac{1}{2}\binom{n}{k}$ monomials of degree k.*

Proof.
A given monomial $x_{i_1} x_{i_2} \ldots x_{i_k}$ of degree k will be part of the ANF if and only if $a_u = 1$ where the support of u (that is to say the set of indices j such that $u_j = 1$ and denoted $supp(u)$) is $\{i_1, i_2, \ldots, i_k\}$. Now we have

$$a_u = f(\overline{0})) \oplus \bigoplus_{j=1}^{k} f(e_{i_j}) \oplus \left(\bigoplus_{l=1}^{k} \bigoplus_{j=1, j \neq l}^{k} f(e_{i_j} \oplus e_{i_l}) \right) \oplus \ldots \oplus f(\bigoplus_{j=1}^{k} e_{i_j}) \tag{3}$$

where $\overline{0} = (0, 0, \ldots, 0)$ and e_i is the n-uple whose only its i-th coordinate is non zero. The right side of Equation 3 has $\sum_{j=1}^{k} \binom{k}{j} = 2^k$ terms. $a_u = 1$ if an odd number of terms are each equal to

1. There are 2^{k-1} such odd configurations. Each of them according to 1 has probability $\frac{1}{2^k}$ to be equal to 1. Whence we have

$$P[a_u = 1] = 2^{k-1} \times \frac{1}{2^k} = \frac{1}{2}$$

Thus the number of monomials of degree k in the ANF will be

$$P[a_u = 1] \times \binom{n}{k} = \frac{1}{2} \times \binom{n}{k}$$

□

We can in fact generalize this results with the following theorem:

Theorem 1 *With the notation used in Proposition 1, the number n_k of monomials of degree k has normal distribution with mean value and variance given by:*

$$E[n_k] = \frac{1}{2}\binom{n}{k} \qquad \text{and} \qquad V[n_k] = \frac{1}{4}\binom{n}{k}$$

Proof.
The proof is straightforward when considering that a_u, for all $u \in \mathbb{F}_2^n$ is a Bernouilli random variable with parameter $\frac{1}{2}$, where $E[a_u] = \frac{1}{2}$ and $V[a_u] = \frac{1}{4}$. Since $n_k = \sum_{wt(u)=k} a_u$, for large enough value of the number of u of weight k, the Central Limit Theorem gives the result. □

This proposition allows to study the randomness properties of a Boolean function. Let us consider a function f used for the feedback of a shift register of length L. If f is constant (its ANF has only one monomial), the output will not be random at all. In the case of the linear feedback (the ANF of f is of degree 1 and has at most n monomials), the randomness properties are limited: the linearity properties are not suppressed, and combinatorial information is easy to get (for details see [18]). Moreover, it is very easy to reconstruct the feedback polynomial with only $2L$ output bits [23]. This is due to the fact that linear functions have very limited randomness properties.

In other words, if we consider $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ such that (*e.g.*) $f(x) = f(y) = 1$, the less random will be the function, the easier will be the extraction of information on x and y.

Example 1 *Let us take $f(x_1, x_2) = x_1 \oplus x_2$. Any $x = (x_1, x_2)$ and $y = (y_1, y_2)$ with $x \neq y$ such that $f(x) = f(y) = 1$ will satisfy $x_1 \oplus y_1 = 1$. This comes from the fact that the values of the truth table are "organized" and not "randomly spread" into this table.*

Proposition 1 gives us the following criterion for Boolean functions suitable for cryptographic applications.

Corollary 1 *A Boolean function used for cryptographic applications and presenting the best trade-off in terms of its cryptographic properties must have a degree as high as possible.*

Proof.
This directly comes from the fact that a n-variable random Boolean function in average has its term of degree n with probability $\frac{1}{2}$ and will contain $\frac{n}{2}$ terms of degree $n-1$. According to the upper bound of the degree [31] of function presenting the best trade-off in terms of correlation immunity, balancedness and nonlinearity we have for a t-correlation immune function:

$$\deg(f(x_1, x_2, \ldots, x_n)) \leq n - t - 1.$$

Imposing to the function to have given properties lowers the algebraic degree. In other words combinatorial structures are introduced while randomness is lessened. In the search for the best possible trade-off, to keep good randomness properties forbidding to get combinatorial information on the function input, the function should have the highest possible degree. □

As a consequence, Boolean functions designed in [11] are more suitable for cryptographic applications than those presented in [22, 32] since these latter have a slightly lower degree. This fact has been confirmed by our tests when considering output sequences produced by nonlinear feedback shift registers. The statistical results are slightly but significantly better for the first one which have been used in the design of COS ciphers [12].

2.2 Characterization of the Walsh Coefficients

The *Walsh Hadamard transform* of a Boolean function f refers to the following transformation:

$$\forall u \in \mathbb{F}_2^n, \quad \widehat{\chi_f}(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+<x,u>}$$

where $< x, u >$ denotes the usual scalar product. A well-known result allows to characterize the correlation immunity of f with the Walsh Hadamard transform:

Proposition 2 *[33] A Boolean function f is t-order correlation immune if and only if*

$$\forall u \in \mathbb{F}_2^n, \quad 1 \leq wt(u) \leq t \qquad \widehat{\chi_f}(u) = 0$$

Moreover f is balanced if and only if $\widehat{\chi_f}(0, 0, \ldots, 0) = 0$. When balanced and t-correlation immune, f is said t-resilient.

Proposition 3 *Let f a random Boolean function over $\mathbb{F}_2^n$ with $n \geq 5$. For all $u \in \mathbb{F}_2^n$, $\widehat{\chi_f}(u)$ is a random variable which has Gaussian distribution with mean value 0 and variance 2^n.*

Proof.
First we can write

$$\begin{aligned} \widehat{\chi_f}(u) &= \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+<x,u>} \\ &= 2^n - 2. \sum_{x \in \mathbb{F}_2^n} (f(x)+ < x, u >) \end{aligned}$$

Since x and $f(x)$ are independent, we can consider $< x, u > +f(x)$ as independent, identically distributed random variables for all x as well. Let us note $Y = \sum_{x \in \mathbb{F}_2^n} (f(x)+ < x, u >)$. For $n > 5$ (that is to say $2^n > 30$), due to the central limit theorem [6], Y has a Gaussian distribution $\mathcal{LG}(E, \sigma^2)$ with

$$\begin{aligned} E[Y] &= 2^n P[f(x)+ < x, u >= 1] = 2^{n-1} \\ (\sigma_Y)^2 &= 2^n P[f(x)+ < x, u >= 1] P[f(x)+ < x, u > \neq 1] = 2^{n-2} \end{aligned}$$

Hence $\widehat{\chi_f}(u)$ has Gaussian distribution with mean value

$$E[\widehat{\chi_f}(u)] = 2^n (1 - 2P[f(x)+ < x, u >= 1]) = 0$$

and variance

$$\sigma^2 = 4.2^n P[f(x)+ < x, u >= 1]P[f(x)+ < x, u >\neq 1] = 2^n$$

□

If Φ denotes the normal distribution function,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$$

and if $p_0 = \Phi(\frac{1}{2^{\frac{n}{2}-1}}) - \frac{1}{2}$ we then can state

Lemma 1

$$P[f \ \ balanced\ ] = p_0.$$

Proof.
In case of balanced Boolean function, we must have $\widehat{\chi_f}(0, 0, \ldots, 0) = 0$ By definition $\widehat{\chi_f}(u) \quad \forall u \in \mathbb{F}_2^n$ is even. We thus can write

$$P[\widehat{\chi_f}(u) = 0] = P[0 < \widehat{\chi_f}(u) < 2]$$

The rest is straightforward to proove with Proposition 3. □

Remark.- This result is an accurate approximation of the "exact" probability for a function to be balanced given by $p = \frac{\binom{2^n}{2^{n-1}}}{2^{2^n}}$. Table 1 compares exact probability with that computed with Theorem 1 for $5 \leq n \leq 19$. Note that computing exact probability p is highly time consuming while computation time is negligible for p_0.

n	p	p_0	n	p	p_0	n	p	p_0
5	0.1399	0.1381	10	0.02493	0.02491	15	0.004408	0.004407
6	0.09935	0.09870	11	0.01763	0.01762	16	0.003117	0.003116
7	0.07039	0.07015	12	0.01247	0.01246	17	0.002204	0.002203
8	0.04982	0.49738	13	0.008815	0.008814	18	0.001558	0.001558
9	0.03524	0.03521	14	0.006233	0.006233	19	0.001102	0.001101

Table 1: Comparison between exact probability p and approximate probability p_0 for a function to be balanced

3 The New Statistical Testing

We are now going to present the different tests we built up to evaluate new statistical properties of symmetric cryptosystems and hash functions. Let us now consider such a system and precise the context we choose. Let be a secret key $K = (k_0, k_1, \ldots, k_{n-1})$. A stream cipher can be seen as follows: each output bits i generated from the secret key K can be expressed by a unique ANF (by means of the Möbius transform defined by Equation 2).

In other words the N-bits output sequence can be described by a family of N Boolean functions

$$(f_t(K))_{0 \leq t < N} = (f_0(K), f_1(K), \ldots, f_{N-1}(K))$$

where $f_i(K)$ denotes the i-th bit produced by the system and modelled as a polynomial in variables k_i (ANF). Each output bit is a Boolean function $f_t : \mathbb{F}_2^n \mapsto \mathbb{F}_2$

Similarly let us represent a block cipher with n-bit key K and working on m-bit blocks. In the same way, but with the different output functions being evaluated on the key space and on the plaintext space $P = (p_0, p_1, \ldots, p_{m-1})$, for a block cipher C, we then have:

$$C = (c_0, c_1, \ldots, c_{m-1}) = (f_0(K, P), f_1(K, P), \ldots, f_{N-1}(K, P))$$

Each of the m ciphertext bits is a Boolean function $f_t : \mathbb{F}_2^{n+m} \mapsto \mathbb{F}_2$

Finally a hash function $H : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ will have its m-bit message digest of block $B = (b_0, b_1, \ldots, b_{n-1})$ represented by:

$$(h_t(B))_{0 \leq t < m} = (h_0(B), h_1(B), \ldots, h_{m-1}(B))$$

In the rest of this paper we will use indifferently the term "output bits" and "output Boolean Functions" (or output ANFs for short) to describe the quantities produced by the cryptosystem we consider. At last we will consider that the different output Boolean functions (or bits) are statistically independant. It is precisely the result stated by previous usual, known tests.

The complete output ANF cannot be computed since it contains in average 2^{n-1} monomials. It would require exponential memory and computing time complexity. For our tests we only focus on the monomials of degree at most 3 and need only to compute the d-truncated ANF, that is to say the partial ANF whose coefficients are effectively computed up to degree 3. From a practical point of view, we use Formula 3 to produce them. As a result, we observe in each ANFs, $\hat{n}_d$ monomials of degree exactly d.

Let us now note H_0^d the statistical hypothesis that the number $\hat{n}_d$ of monomials of degree exactly d is distributed according to the Theorem 1. In other words the cryptosystem pass our tests and thus exhibit no particuliar structural, statistical bias for the aspect we consider when satisfying this hypothesis.

We will not recall basic probability and statistics theory. We suppose the reader familiar with them (for detailed presentation see [6] and [21, Chap 5.4]).

3.1 The Affine Constant Test

Out hypothesis is then denoted H_0^0. According to Theorem 1, the probability for the affine constant a_0 to be represented in each of the output ANFs is $p = \frac{1}{2}$. Equivalently it means that the number of output Boolean functions having $a_0 = 1$ in their ANF has normal distribution $\mathcal{N}(\frac{N}{2}, \frac{\sqrt{N}}{2})$ where N is the total number of output ANFs.

If X_S, the number of times $a_0 = 1$, is the statistic we consider over the sample output S of N ANFs, we can now describe the following two-sided test, called the *Affine Constant Test*:

1. Compute X_S over S.
2. Let us fix a significance level α (*i.e.* probability of rejecting H_0^0 when it is true) and choose a threshold x_α so that for a statistic X of normal standard distribution we have
$$P[X > x_\alpha] = P[X < x_\alpha] = \frac{\alpha}{2}$$
3. If the value $\hat{X}_S = \frac{X_S - \frac{N}{2}}{\frac{\sqrt{N}}{2}} > x_\alpha$ or if $\hat{X}_S < -x_\alpha$ then H_0^0 is rejected (the system fails the test) otherwise H_0^0 is kept (the system passes the test).

In our experiments, we considered $\alpha = 0.05, 0.01$ and 0.001.

3.2 The d-monomial Tests

We are now considering the monomials of degree exactly d in the output ANFs. Our testing is now denoted H_0^d.

With the notation of Theorem 1, the number of monomials of degree d in a Random Boolean Function ANF (RBFANF) is a random variable which is $\mathcal{N}(\frac{1}{2}\binom{n}{d}, \frac{1}{2}\sqrt{\binom{n}{d}})$ distributed. We now consider two *goodness-of-fit*, one-sided tests between the expected frequencies (denoted n_d) and those (denoted $\hat{n}_d$) we observe for the considered cryptosystem.

The first test, T_1^d consider every different ANF and thus has a rather local scope by giving more weight to very weak output ANFs. The second one, T_2^d, groups the N RBFANFs according to a few numbers of sets or classes. So to summarize, we will use the χ^2 distribution with ν degrees of freedom by considering the sum of the ν squared, independent random variables $\frac{(n_d^i - \hat{n}_d^i)}{\sqrt{n_d^i}}$ $(i \leq \nu)$ which have by definition standard normal distribution.

In T_1^d we have $\nu = N - 1$ (*i.e.* the number of output ANFs) while for T_2^d we choose $2 \leq \nu \leq 9$.

1. Compute for each of the ν random variables n_d^i and $\hat{n}_d^i$ (n_d^i is given by applying Theorem 1).
2. Let us fix a significance level α and a threshold value x_α (computed directly from the cumulative density function of the χ^2 distribution) so that for a statistic X over a random sample we would have $P[X > x_\alpha] = \alpha$ (when X follows a χ^2 distribution with ν degrees of freedom).
3. Compute the statistics D^2 given by

$$D^2 = \sum_{i=1}^{\nu} \frac{(n_d^i - \hat{n}_d^i)^2}{n_d^i}$$

4. If $D^2 > x_\alpha$ then we must reject H_0^d (the system fails the test and thus there is a statistical bias) otherwise we keep H_0^d (the system does not present any significative bias).

In our experiments, we considered $\alpha = 0.05, 0.01$ and 0.001.

Test T_2^d is intended to describe the considered cryptosystem from a global point of view. In particular it aims at verifying if local biases (detected with T_1^d) are still really significative at a more global level. Instead of dealing with the observed frequencies $\hat{n}_d^i$ of d-monomials for each of the N output ANFs we rather are interested with the number of output ANFs whose number $\hat{n}_d$ belongs to a given, predefined[1] intervall $[a, b[$. The expected frequency for each class is computed from Theorem 1 by applying basic probability results.

4 Testing Results on Symmetric Systems and Hash Functions

4.1 Stream Ciphers

We will here mainly focus on two stream ciphers that have been proposed for the NESSIE Open Call for Cryptographic Primitives: Lili-128 and Snow. For information on NESSIE Project and these two algorithms see [25]. Other stream ciphers have been tested or are currently under testing. Table 2 summarizes results for a few of them. For complete, detailed results see [7]. We considered the first $N = 6016$ output bits in our experiments. It is worth noticing that:

[1] In fact according to the probability theory [6], the only constraint is that the expected frequency for each class, given by $N.P[a < n_d < b]$ must be greater than 5; otherwise the intervalls can be freely chosen

- All the tested stream ciphers pass the Affine Constant test except Lili-128.
- Lili-128 exhibits extremely strong biases. Table 3 presents the results for this stream cipher. These biases have been analyzed and exploited for an operational cryptanalysis in [10] (see Section 5 concerning the cryptanalysis part).
- Snow exhibits strong biases too but only when considering global statistical behavior. Unfortunately these biases allowed us to design a complete, operationnal cryptanalysis of Snow [9] (see Section 5 concerning the cryptanalysis part).
- We can give the following interesting observations based on the comparison of the tests convergence (that is to say the distance between the estimator and the threshold value; for details see [19]). The stream ciphers of Table 2 can be sorted according to their relative "random" quality. For example when considering results of test T_1^1 (1-monomials), which is the most interesting, we have the following ordering ($\succeq$ means "better than"):

$$\text{Bgml} \succeq RC4 \succeq \text{Snow} \succeq\succeq \text{Lili-128}$$

	T_1^1	T_1^2	T_2^1	T_2^2
Lili-128	**fail**	**fail**	**fail**	**fail**
Snow	pass	pass	**fail**	**fail**
RC4 [27]	pass	pass	pass	pass
Bgml [25]	pass	pass	pass	pass

Table 2: Stream Ciphers: Tests Results (significance level $\alpha = 0.05, 0.01, 0.001$)

	T_1^1	T_1^2	T_2^1	T_2^2
D^2	39,344.03	400,839.93	667729.02	1,028,048.45
$\chi^2_{0.05}$	6196.27			
$\chi^2_{0.01}$	6272.76			
$\chi^2_{0.001}$	6349.15			

Table 3: Lili128: Experimental results for tests T_1^d and T_2^d.

4.2 Block Ciphers

We mainly focus on the DES [13] and the AES [1]. For block ciphers we considered both the encryption ANFs and the decryption ANFs. Since the output ANF involves both plaintext variables and key variables, tests T_2^d ($d = 1, 2$) have been replaced by tests T_1^d relatively to:

- the number n_1 of plaintext variables from one side and of key variables from the other side (denoted respectively $T_1^1|p$ and $T_1^1|k$).
- the number n_2 of 2-monomials respectively involving plaintext/plaintext variables, key/key variables and plaintext/key variables (tests denoted respectively $T_1^1|pp$, $T_1^1|kk$ and $T_1^1|pk$.

4.2.1 The DES

Table 4 summarizes the results of the different tests. Table 5 gives detailed experimental results

	T_1^1	T_1^2	$T_1^1\|p$	$T_1^1\|k$	$T_1^1\|pp$	$T_1^1\|kk$	$T_1^1\|pk$
Encryption with IP	pass	**fail**	pass	pass	**fail**	pass	**fail**
Encryption without IP	pass	**fail**	pass	pass	**fail**	pass	**fail**
Decryption with IP	pass	**fail**	pass	pass	**fail**	pass	**fail**
Decryption without IP	pass	**fail**	pass	pass	**fail**	pass	**fail**

Table 4: DES: Tests Results (significance level $\alpha = 0.05, 0.01, 0.001$)

of the estimator D^2 with 63 degrees of freedom. These values are to be compared to theoretical values $\chi^2 = 82.52$ for $\alpha = 0.05$, $\chi^2 = 92.01$ for $\alpha = 0.01$ and $\chi^2 = 103.44$ for $\alpha = 0.001$. Complete intermediate data will be found in [7]. It is worth noticing that:

- DES passes the Affine Constant Test in all modes and all significance level.
- The overall statistical quality is better for encryption than for decryption.
- The initial permutation IP improves the overall statistical quality. Nevertheless IP is usually discarded by cryptology community when considering its cryptanalysis.
- When considering 2-monomials, DES exhibit very strong biases (except for key/key monomials).

4.2.2 The AES

We will focus on the algorithm working on 128-bit blocks and with 128-bit secret key. Results for other versions can be found in [7] as well as complete intermediate data. Table 6 gives detailed experimental results of the estimator D^2 with 127 degrees of freedom. These values are to be compared to theoretical values $\chi^2 = 159.59$ for $\alpha = 0.05$, $\chi^2 = 166.27$ for $\alpha = 0.01$ and $\chi^2 = 180.61$ for $\alpha = 0.001$. It is worth noticing that:

- AES passes the Affine Constant Test in all modes and all significance level.
- Overall statistical quality of AES (128, 128) is good except for plaintext/plaintext 2-monomials for which AES fails the test whatever may be the significance level. These biases have been recently exploited to greatly improve the cryptanalysis of AES (see Section 5).
- Encryption and decryption exhibits quite the same overall statistical properties.

4.3 Hash Functions

We tested the following hash functions: SHA-0 [15], SHA-1 [16], Ripemd160 [4], MD4 [28], MD5 [29], Ripe-MD [3] and Haval [34] (for this latter we tested all the different versions). Extensively detailed numerical results (due to lack of space) are only available in [7].

All the tested hash functions have passed the tests whatever may be the significance level. However we can once again give the following interesting observations based on the comparison of the tests convergence.

	T_1^1	T_1^2	$T_1^1\|p$	$T_1^1\|k$	$T_1^1\|pp$	$T_1^1\|kk$	$T_1^1\|pk$
Encr. + IP	40.44	2662.27	54.12	29.12	1909.14	36.46	1794.70
Encr. - IP	46.51	2693.27	54.12	41.17	1909.14	34.69	1817.14
Decr. + IP	41.03	3017.51	65.53	29.67	2287.92	36.46	1889.92
Decr. - IP	51.00	3005.31	65.53	41.17	2287.92	34.69	1880.62

Table 5: DES: Values of Estimator D^2

	T_1^1	T_1^2	$T_1^1\|p$	$T_1^1\|k$	$T_1^1\|pp$	$T_1^1\|kk$	$T_1^1\|pk$
Encryption	59.61	157.91	57.84	61.51	**415.20**	72.34	62.39
Decryption	67.38	156.03	67.21	70.70	**412.87**	60.11	47.27

Table 6: AES (128, 128): Values of Estimator D^2

- The different hash functions can be sorted according to their relative "random" quality. For example when considering results of test T_1^1 (1-monomials), which is the most interesting, we have the following ordering ($\succeq$ means "better than"):
  - 160-bit Message Digest: SHA-1 $\succeq$ (5, 160)-haval $\succeq$ Ripemd160 $\succeq$ (4, 160)-haval $\succeq$ (3, 160)-haval $\succeq$ SHA-0.
  - 128-bit Message Digest: (5,128)-haval $\succeq$ Ripe-MD $\succeq$ MD5 $\succeq$ (4,128)-haval $\succeq$ (3,128)-haval $\succeq$ MD4.
- SHA-1 has indeed better statistical properties than SHA-0, especially when considering the degree 1. The inclusion of the 1-bit rotation in the block expansion from 16 to 80 words really improved the randomness properties of the hash function.
- For the Haval family, the random quality increases with the number of rounds.

Table 7 presents the results of the tests T_1^d and T_2^d for $d = 1, 2$ and for the 160-bit message digest hash functions (significance level $\alpha = 0.05$; let us recall that passing the tests for significance level α imply passing the test for $\alpha' < \alpha$ since $\chi^2_{\alpha'} > \chi^2_{\alpha}$). All other results will be found in [7].

Hash Functions	T_1^1		T_1^2		T_2^1		T_2^2	
	D^2	χ^2	D^2	χ^2	D^2	χ^2	D^2	χ^2
SHA-1	76.87	189.52	70.89	189.52	0.04	5.99	0.42	5.99
(5,160)-haval	76.34		79.76		0.17		2.02	
Ripemd160	77.51		66.72		5.24		2.66	
(4,160)-haval	83.52		74.18		1.77		3.51	
(3,160)-haval	83.79		64.28		1.05		5.50	
SHA-0	97.08		74.50		3.26		0.42	

Table 7: Experimental results for tests T_1^d and T_2^d ($d = 1, 2$, $\alpha = 0.05$).

5 Conclusion

This paper presented a new statistical testing of symmetric ciphers and hash functions. Where previous known tests did not exhibit particuliar bias, these new tests reveal structural, statistical

biases for DES, AES and Lili-128. Other cryptosystems are currently tested and may present unsuspected biases.

These tests are still rather quantitative tests but allow to detect possible structural weaknesses in the output ANFs. Current research focuses on more qualitative test involving factorial experiments. It should provide necessary information to greatly improve previous cryptanalytic techniques. But it has already been possible to design a completely new, deterministic, OPERATIONNAL (both in terms of computing time and of number of required output bits) cryptanalysis of Lili-128 and Snow. The biases we have detected with this new testing, have been successfully converted in purely combinatorial properties allowing a deterministic cryptanalysis. For the AES, the biases have been recently exploited to design a new statistical, combinatorial cryptanalysis of AES [8]. In both cases, the cryptosystems are mainly modelled by special combinatorial designs.

References

[1] *http://www.nist.gov/aes/*

[2] H. Beker, F. Piper, *Cipher Systems: The Protection of Communications*, John Wiley & Sons, New York, 1982.

[3] A. Bosselaers, B. Preenel editors, *Intregrity Primitives for Secure Information Systems: Final Report of RACE Integrity Primitives Evaluation RIPE-RACE 1040*, Lecture Notes in Computer Science 1007, Springer Verlag, 1995.

[4] H. Dobbertin, A. Bosselaers, B. Preenel, RIPEMD-160: a Strengthened Version of RIPEMD. In. *D. Gollman ed., Fast Software Encryption, Third International Workshop*, Lecture Notes in Computer Science 1039, pp 71–82, Springer-Verlag, 1996.

[5] E.D. Erdmann, *Empirical Tests of Binary Keystreams*, Master's thesis, Department of Mathematics, Royal Holloway and Bedford New College, University of London, 1992.

[6] W. Feller, *An Introduction to Probability Theory*, Wiley, 1966.

[7] *http://www-rocq.inria.fr/codes/index.html* (the complete materials will be avalaible by September 2002).

[8] E.Filiol, *A New Cryptanalysis of Block Ciphers: the AES Case*, Private Report, 2002.

[9] E. Filiol, *Operational Cryptanalysis of Snow*, Private Report, 2002.

[10] E.Filiol, Strong Weaknesses in Lili-128 Stream Cipher, Private Report, 2002.

[11] E. Filiol, C. Fontaine, Highly Nonlinear Balanced Boolean Functions with a Good Correlation-Immunity, *Advances in Cryptology - EUROCRYPT'98*, Lecture Notes in Computer Sciences 1403, Springer Verlag, 1998.

[12] E. Filiol, C. Fontaine A new Block Cipher Design: COS Ciphers, *Proceedings of the International Symposium on Information Theory 2001*, p. 134, Washington, 2001.

[13] FIPS 46, *Data Encryption Standard*, Federal Information Processing Standards Publication 140-1, US Dept of Commerce/NIST, National Technical Information Service, Springfield, Virginia, 1977.

[14] FIPS 140-1, *Security Requirements for Cryptographic Modules*, Federal Information Processing Standards Publication 140-1, US Dept of Commerce/NIST, National Technical Information Service, Springfield, Virginia, 1994.

[15] FIPS 180, *Secure Hash Standard*, Federal Information Processing Standards Publication 180, US Dept of Commerce/NIST, National Technical Information Service, Springfield, Virginia, 1993.

[16] FIPS 180-1, *Secure Hash Standard*, Federal Information Processing Standards Publication 180-1, US Dept of Commerce/NIST, National Technical Information Service, Springfield, Virginia, 1995.

[17] H. Gustafson, E. Dawson, J. Dj. Golic, Randomness Measures Related to Subset Occurence, In: *Cryptography: Policy and Algorithms, International Conference*, Lecture Notes in Computer Science 1029, pp 132–143, Springer Verlag, 1996.

[18] S.W. Golomb, *Shift Register Sequences*, Aegean Park Press, 1982.

[19] R.V. Hogg, E.A. Tanis, *Probability and Statistical Inference*, MacMillan Publishing, New York, 3rd edition, 1988.

[20] D.E. Knuth *The Art of Computer Programming*, Volume 2, Addison Wesley, Reading, Massachussets, 2nd edition, 1981.

[21] A.J. Menezes, P.C. Van Oorschot, S.A. Vanstone, *Handbook of Applied Cryptography*. CRC Press, 1997.

[22] S. Maitra, P. Sarkar, Construction of Nonlinear Boolean Functions with Important Cryptographic Properties, *Advances in Cryptology - EUROCRYPT'00*, Lecture Notes in Computer Sciences 1807, Springer Verlag, 2000.

[23] J.L. Massey, Shift-Register Synthesis and BCH Decoding, *IEEE Transactions on Information Theory*, Vol. IT-15, pp 122–127, 1969.

[24] U. Maurer, A Universal Statistical Test for Random Bit Generators, *Journal of Cryptology, 5* pp 89-105, 1992.

[25] *http://www.cryptonessie.org*

[26] D. Olejár, M. Stanek, On Cryptographic Properties of Random Boolean Functions, *Electronic Journal of Universal Computer Science, Vol. 4, Issue 8*, 1998.

[27] B. Schneier, *Applied Cryptography: Protocols, Algorithms and Source Code in C*, Wilew et Sons, 2nd edition, 1996.

[28] R.L. Rivest, The MD4 Message Digest Algorithm, *Advances in Cryptology - CRYPT0'90*, Lecture Notes in Computer Sciences 537, Springer Verlag, 1991.

[29] R.L. Rivest, The MD5 Message Digest Algorithm, Internet Request for Comment 1321, April 1992, (presented at the Rump Session of Crypto'91). Lecture Notes in Computer Sciences 537, Springer Verlag, 1991.

[30] C.E. Shannon, Communication Theory of Secrecy Systems, *Bell System Technical Journal, 28*, pp 656–715, 1949.

[31] T. Siegenthaler, Correlation Immunity of Nonlinear Combining Functions for Cryptographic Applications, *IEEE Transactions on Information Theory*, Vol. IT 35, pp 776–780, 1984.

[32] Y. Tarannikov, on Resilient Boolean Functions with Maximal Possible Nonlinearity, *Proceedings of the First International Conference in India - INDOCRYPT'00*, Lecture Notes in Computer Science 1977, Springer Verlag, 2000.

[33] G. Xiao, J.L. Massey, A Spectral Characterization of Correlation Immune Functions, *IEEE Transactions on Information Theory*, Vol. IT-34, pp 569–571, 1988.

[34] Y. Zheng, J. Pieprzyk, J. Seberry, HAVAL - A One-way Hashing Algorithm with Variable Length of Output, *Advances in Cryptology - AUSCRYPT'92*, Lecture Notes in Computer Sciences 718, pp 83–104, Springer Verlag, 1993.