Resistance Against General Iterated Attacks

Serge Vaudenay

Ecole Normale Supérieure — CNRS
Serge.Vaudenay@ens.fr

Abstract. In this paper we study the resistance of a block cipher against any general iterated attack. This class of attacks includes differential and linear cryptanalysis. We prove that we can upper bound the complexity of the attack by using Vaudenay's decorrelation technique. Our main theorem enables to prove the security of some recently proposed block ciphers COCONUT98 and PEANUT98.

Since public-key cryptography has been discovered in the late 70s, *proving* the security of cryptographic protocols has been a challenging problem. Recently, the random oracle model and the generic algorithm techniques have introduced new tools for validating cryptographic algorithms. Although much older, the area of symmetric cryptography did not get so many tools.

In the early 90s, Biham and Shamir [2] introduced the notion of differential cryptanalysis and Matsui [7, 8] introduced the notion of linear cryptanalysis, which was a quite general model of attack. Since then many authors tried to formalize these attacks and study their complexity in order to prove the security of block ciphers against it. Earlier work, initiated by Nyberg [9] was based on algebraic techniques.

Recently, Vaudenay adapted Carter and Wegman's combinatoric notion of universal functions [3, 16] in context of encryption and formalized the notion of decorrelation bias [13]. This measurement enables to quantify the security of block ciphers against several classes of attacks. In [13], several real-life block ciphers have been proposed, namely COCONUT98 and PEANUT98. Their decorrelation bias have been measured, and the security against differential and linear cryptanalysis has been proved.

In this paper, we generalize these results in a uniform approach. We introduce the notion of *iterated attack of order d* and we prove

how the decorrelation bias can measure the security against any of it. Differential and linear cryptanalysis happen to be included in this class of attacks.

This paper is organized as follows. First we recall the previous results in decorrelation theory which are interesting for our purpose in Section 1. Our contribution starts on page 6 in Section 2. We define the notion of iterated attack of given order. We prove that decorrelation of order d is not sufficient to thwart all iterated attacks of order d with a counterexample. We then show how decorrelation of order $2d$ gives an upper bound on the efficiency of any iterated attacks of order d. We show how to use this result for a practical block cipher (namely, PEANUT98). Finally, in Section 3 we investigate how to use the same techniques for combining several cryptanalysis all together.

1 Previous Work

We briefly recall some of Vaudenay's definitions and results about decorrelation which are taken from [13].

1.1 Decorrelation Theory

In this setup, a block cipher is considered as a random permutation C over a message-block space $\mathcal{M}$. (Here the randomness comes from the random choice of the secret key.) The efficiency of a cryptanalysis can be measured by the average complexity of the algorithm over the distribution of the permutation (*i.e.* of the secret key).

Definition 1. *Given a random function F from a given set $\mathcal{M}_1$ to a given set $\mathcal{M}_2$ and an integer d, we define the* d-wise distribution matrix $[F]^d$ *of F as a $\mathcal{M}_1^d \times \mathcal{M}_2^d$-matrix where the (x, y)-entry of $[F]^d$ corresponding to the multi-points $x = (x_1, \ldots, x_d) \in \mathcal{M}_1^d$ and $y = (y_1, \ldots, y_d) \in \mathcal{M}_2^d$ is defined as the probability that we have $F(x_i) = y_i$ for $i = 1, \ldots, d$.*

Basically, each row of the d-wise distribution matrix corresponds to the distribution of the d-tuple $(F(x_1), \ldots, F(x_d))$ where $(x_1, \ldots, x_d)$ corresponds to the index of the row.

In this paper, we consider the following matrix norm over $\mathbf{R}^{\mathcal{M}^d \times \mathcal{M}^d}$ defined by

$$||A|| = \max_x \sum_y |A_{x,y}|$$

for any matrix A.

Definition 2. *Given two random functions F and G from a given set $\mathcal{M}_1$ to a given set $\mathcal{M}_2$, an integer d, we call $||[F]^d - [G]^d||$ the* d-wise decorrelation distance *between F and G.*

A decorrelation distance of zero means that for any multi-point $x = (x_1, \ldots, x_d)$ the multi-points $(F(x_1), \ldots, F(x_d))$ and $(G(x_1), \ldots, G(x_d))$ have the same distribution, so that F and G have the same *decorrelation*.

Random permutations C are compared to a uniformly distributed permutations C^* (which will be called *perfect cipher*, and the decorrelation of which will be said to be *perfect*). For instance, saying that a *cipher* C on $\mathcal{M}$ has a perfect pairwise decorrelation means that for any $x_1 \neq x_2$, the random variable $(C(x_1), C(x_2))$ is uniformly distributed among all the (y_1, y_2) pairs such that $y_1 \neq y_2$.

Definition 3. *Let C be a random permutation over $\mathcal{M}$. We call* d-wise decorrelation bias *and we denote* $\mathrm{Dec}^d(C)$ *the quantity $||[C]^d - [C^*]^d||$ where C^* is uniformly distributed.*

This notion is fairly similar to the notion of universal functions which was been introduced by Carter and Wegman [3, 16].

The matrix norm property (*i.e.* $||A \times B|| \leq ||A||.||B||$) implies

$$\mathrm{Dec}^d(C_1 \circ C_2) \leq \mathrm{Dec}^d(C_1).\mathrm{Dec}^d(C_2).$$

Thus we can built ciphers with arbitrarily small decorrelation bias by iterating a simple cipher as long as its own decorrelation bias is smaller than 1. The security results show that when the decorrelation bias is small, then the complexity of the attack is high.

1.2 Security Model

In the Luby-Rackoff model, an attacker is an infinitely powerful Turing machine $\mathcal{A}^{\mathcal{O}}$ which has access to an oracle $\mathcal{O}$ whose aim is to distinguish a cipher C from the Perfect Cipher C^* by querying the

oracle which implements either cipher, and with a limited number d of inputs (see [6]). The oracle $\mathcal{O}$ either implements C or C^*, and that the attacker must finally answer 0 ("reject") or 1 ("accept"). We measure the ability to distinguish C from C^* by the advantage $\mathrm{Adv}_{\mathcal{A}}(C) = |p - p^*|$ where p (resp. p^*) is the probability of answering 1 if $\mathcal{O}$ implements C (resp. C^*). In this paper we focus on non-adaptive attacks *i.e.* on distinguishers illustrated on Fig. 1: here no x_i queried to the oracle depends on some previous answers y_j.

Input: an oracle which implements a permutation c
1. calculate some messages $X = (X_1, \ldots, X_d)$
2. get $Y = (c(X_1), \ldots, c(X_d))$
3. depending on X and Y, output 0 or 1

Fig. 1. A d-Limited Non-Adaptive Distinguisher.

Theorem 4. *Let d be an integer. Let C be a cipher. The best d-limited non-adaptive distinguisher $\mathcal{A}$ for C is such that*

$$\mathrm{Adv}_{\mathcal{A}}(C) = \frac{1}{2}\mathrm{Dec}^d(C).$$

1.3 Differential and Linear Cryptanalysis

In this section we assume that $\mathcal{M} = \mathrm{GF}(2^m)$. The inner dot product $a \cdot b$ in $\mathrm{GF}(2^m)$ is the parity of the bitwise *and* of a and b.

We call basic differential (resp. linear) cryptanalysis the distinguisher which is characterized by a pair $(a, b) \in \mathcal{M}^2$ and which is depicted on Fig. 2 (resp. Fig. 3). Linear cryptanalysis also needs an *acceptance set* B.

Theorem 5. *Let C be a cipher on a group $\mathcal{M}$ of order M. For any basic differential distinguisher (depicted on Fig. 2) of complexity n, we have*

$$\mathrm{Adv}_{\mathrm{Fig.2}}(C) \leq \frac{n}{2^m - 1} + \frac{n}{2}\mathrm{Dec}^2(C).$$

Input: a cipher c, a complexity n, a characteristic (a, b)
1. for i from 1 to n do
   (a) pick uniformly a random X and query for $c(X)$ and $c(X + a)$
   (b) if $c(X + a) = c(X) + b$, stop and output 1
2. output 0

Fig. 2. Differential Distinguisher.

Input: a cipher c, a complexity n, a characteristic (a, b), a set B
1. initialize the counter value u to zero
2. for i from 1 to n do
   (a) pick a random X with a uniform distribution and query for $c(X)$
   (b) if $X \cdot a = c(X) \cdot b$, increment the counter u
3. if $u \in B$, output 1, otherwise output 0

Fig. 3. Linear Distinguisher.

Theorem 6. *Let C be a cipher on $\mathcal{M} = \{0,1\}^m$. For any linear distinguisher (depicted on Fig. 3) we have*

$$\lim_{n \to +\infty} \frac{\mathrm{Adv}_{\mathrm{Fig.3}}(C)}{n^{\frac{1}{3}}} \leq 9.3 \left(\frac{1}{2^m - 1} + 2\mathrm{Dec}^2(C) \right)^{\frac{1}{3}}.$$

This asymptotic result comes from approximation to the normal law by the Central Limit Theorem, which is correct whenever n is not too small (*i.e.* $n > 30$). This result is thus actually valid for any practical n. So, if the pairwise decorrelation bias has the order of 2^{-m}, linear distinguishers does not work against C, but with a complexity in the scale of 2^m.

1.4 Some Constructions

In [13], two real-life block ciphers have been proposed: COCONUT98 and PEANUT98. They come from the general COCONUT and PEANUT family constructions respectively.

A cipher in the COCONUT family is characterized by some parameters (m, p) where m is the message-block length and p is an irreducible polynomial of degree m in GF(2). The COCONUT98 Cipher corresponds to the parameters $m = 64$ and $p = x^{64} + x^{11} + x^2 + x + 1$.

From the construction, any of COCONUT ciphers has a perfect pairwise decorrelation. Therefore from Theorems 5 and 6 no differential or linear distinguisher can be efficient.

A cipher in the PEANUT family is characterized by some parameters (m, r, d, p) where m is the message-block length, r is the number of rounds (actually, a PEANUT cipher is an r-round Feistel cipher [4]), d is the order of constructed decorrelation, and p is a prime number greater than $2^{\frac{m}{2}}$. The PEANUT98 Cipher corresponds to $m = 64$, $r = 9$, $d = 2$ and $p = 2^{32} + 15$. We have the following result.

Theorem 7. *Let C be a cipher in the PEANUT family with parameters (m, r, d, p). We have*

$$\mathrm{Dec}^d(C) \leq \left(\left(1 + 2\left(p^d 2^{-\frac{md}{2}} - 1\right)\right)^3 - 1 + \frac{2d^2}{2^{\frac{m}{2}}} \right)^{\lfloor \frac{r}{3} \rfloor}.$$

When $p \approx 2^{\frac{m}{2}}$ we can approximate

$$\mathrm{Dec}^d(C) \approx \left(\frac{6d\left(p - 2^{\frac{m}{2}}\right) + 2d^2}{2^{\frac{m}{2}}} \right)^{\lfloor \frac{r}{3} \rfloor}.$$

Hence for the PEANUT98 Cipher we have $\mathrm{Dec}^2(C) \leq 2^{-76}$. Therefore from Theorems 5 and 6 no differential or linear distinguisher can be efficient.

2 Iterated Attacks of Order *d*

In this section we introduce the notion of *iterated attack*.

2.1 Definition

Theorems 5 and 6 suggest that we try to generalize them to attacks in the model depicted on Fig. 4. In this model, we iterate a d-limited non-adaptive attack $\mathcal{T}$. We assume that this attack obtains a sample (X, Y) with $X = (X_1, \ldots, X_d)$ and $Y = (Y_1, \ldots, Y_d)$ such that $y_i = c(X_i)$ for a given distribution of X. Thus, we can think of a known plaintext attack where X has a fixed distribution (*e.g.* a

uniform distribution) or of a chosen plaintext attack where X has a given distribution (*e.g.* in differential cryptanalysis, $X = (X_1, X_1+a)$ where X_1 has a uniform distribution). The result of the attack depends on the result of all iterated ones in a way characterized by a set A.

Input: a cipher c, a complexity n, a distribution on X, a test $\mathcal{T}$, an acceptance set A

1. for i from 1 to n do
   (a) get a new $X = (X_1, \ldots, X_d)$
   (b) get $Y = (c(X_1), \ldots, c(X_d))$
   (c) set $T_i = 0$ or 1 with an expected value $\mathcal{T}(X, Y)$
2. if $(T_1, \ldots, T_n) \in A$ output 1 otherwise output 0

Fig. 4. Iterated Attack of Order d.

For instance, differential cryptanalysis is an iterated attack of order $d = 2$ with $A = \{0,1\}^n \backslash \{(0, \ldots, 0)\}$. Similarly, linear cryptanalysis is an iterated attack of order $d = 1$ with

$$A = \{(t_1, \ldots, t_n); t_1 + \ldots + t_n \in B\}$$

for a given set B.

2.2 A Counterexample

It is tenting to believe that a cipher resists to this model of attacks once it has a small d-wise decorrelation bias. This is wrong as the following example shows. Let C be a cipher with a perfect d-wise decorrelation. We assume that an instance c of C is totally defined by d (x_i, y_i) points so that C is uniformly distributed in a set of $K = M(M-1)\ldots(M-d+1)$ permutations denoted $c_1, \ldots, c_K$. From $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ we can define $I(x, y)$ as the unique index k such that $c_k(x_i) = y_i$ for $i = 1, \ldots, d$. We let

$$\mathcal{T}(x, y) = \begin{cases} 1 \text{ if } I(x, y) \equiv 0 \bmod \mu \\ 0 \text{ otherwise} \end{cases}$$

for a given modulus $\mu = n/a$ and

$$\mathcal{A} = \{0,1\}^n \backslash \{(0, \ldots, 0)\}.$$

If we feed this attack with C or C^*, we have

$$p \approx \frac{1}{\mu} = \frac{a}{n} \text{ and } p^* \approx 1 - \left(1 - \frac{1}{\mu}\right)^n \approx 1 - e^{-a}$$

for $a << n$. Thus Adv can be large even with a relatively large n. This problem actually comes from the fact that the tests $\mathcal{T}$ provide a same expected result for C and C^* but a totally different standard deviation.

2.3 Security Result

We can however prove the security when the cipher has a good decorrelation to the order $2d$.

Theorem 8. *Let C be a cipher on a message space of size M such that $\mathrm{Dec}^{2d}(C) \leq \epsilon$ for some given $d \leq M/2$. For any iterated attack (depicted on Fig. 4) of order d such that the obtained plaintexts are independent, we have*

$$\mathrm{Adv}_{\mathrm{Fig.4}}(C) \leq 3\left(\left(2\delta + \frac{5d^2}{2M} + \frac{3\epsilon}{2}\right)n^2\right)^{\frac{1}{3}} + \frac{n\epsilon}{2}$$

where δ is the probability that for two independent X and X' there exists i and j such that $X_i = X'_j$.

For instance, if the distribution of X is uniform, we have $\delta \leq \frac{d^2}{2M}$.

Proof. Let Z (resp. Z^*) be the probability that the test accepts $(X, C(X))$ (resp. $(X, C^*(X))$), *i.e.*

$$Z = E_X(\mathcal{T}(X, C(X))).$$

Let p (resp. p^*) be the probability that the attack accepts, *i.e.*

$$p = \Pr_C[(T_1, \ldots, T_n) \in A].$$

Since the T_is are independent and with the same expected value Z which only depends on C, we have

$$p = E_C\left(\sum_{(t_1,\ldots,t_n)\in A} Z^{t_1+\ldots+t_n}(1-Z)^{n-(t_1+\ldots+t_n)}\right).$$

We thus have $p = E(f(Z))$ where $f(z)$ is a polynomial of degree at most n with values in $[0,1]$ for any $z \in [0,1]$ entries and with the form $f(z) = \sum a_i z^{b_i}(1-z)^{n-b_i}$. It is straightforward that $|f'(z)| \leq n$ for any $z \in [0,1]$. Thus we have $|f(z) - f(z^*)| \leq n|z - z^*|$.

The crucial point in the proof is in proving that $|Z - Z^*|$ is small within a high probability. For this, we need $|E(Z) - E(Z^*)|$ and $|V(Z) - V(Z^*)|$ to be both small.

From Theorem 4 we know that $|E(Z) - E(Z^*)| \leq \frac{\epsilon}{2}$. We note that Z^2 corresponds to a another test but with $2d$ entries, hence we have $|E(Z^2) - E((Z^*)^2)| \leq \frac{\epsilon}{2}$. Hence $|V(Z) - V(Z^*)| \leq \frac{3}{2}\epsilon$. Now from the Tchebichev's Inequality we have

$$\Pr[|Z - E(Z)| > \lambda] \leq \frac{V(Z)}{\lambda^2}.$$

Hence we have

$$|p - p^*| \leq \frac{2V(Z^*) + \frac{3}{2}\epsilon}{\lambda^2} + n\left(\frac{\epsilon}{2} + 2\lambda\right)$$

so, with $\lambda = \left(\frac{2V(Z^*)+\frac{3}{2}\epsilon}{n}\right)^{\frac{1}{3}}$ we have

$$|p - p^*| \leq 3\left(\left(2V(Z^*) + \frac{3\epsilon}{2}\right)n^2\right)^{\frac{1}{3}} + \frac{n\epsilon}{2}.$$

Now we have

$$\begin{aligned} V(Z^*) &= \sum_{\substack{(x,y)\in A \\ (x',y')\in A}} \Pr_X[x]\Pr_X[x']\left(\Pr_{C^*}\begin{bmatrix} x \to y \\ x' \to y' \end{bmatrix} - \Pr_{C^*}[x \to y]\Pr_{C^*}[x' \to y']\right) \\ &\leq \frac{1}{2}\sum_{\substack{x,y \\ x',y'}} \Pr_X[x]\Pr_X[x']\left|\Pr_{C^*}\begin{bmatrix} x \to y \\ x' \to y' \end{bmatrix} - \Pr_{C^*}[x \to y]\Pr_{C^*}[x' \to y']\right|. \end{aligned}$$

The sum over all x and x' entries with colliding entries (*i.e.* with some $x_i = x'_j$) is less than δ. The sum over all y and y' entries with colliding entries and no colliding x and x' is less than $d^2/4M$. The sum over all no colliding x and x' and no colliding y and y' is equal to

$$\frac{1-\delta}{2}\left(1 - \frac{M(M-1)\ldots(M-2d+1)}{M^2(M-1)^2\ldots(M-d+1)^2}\right)$$

which is less than $\frac{d^2}{2(M-d)}$. Thus we have $V(Z^*) \leq \delta + \frac{d^2}{4M} + \frac{d^2}{2(M-d)}$ which is less than $\delta + \frac{5d^2}{4M}$ when $2d \leq M$. □

This theorem proves that we need $n = \Omega(1/\sqrt{\epsilon})$ or $n = \Omega(\sqrt{M})$ to have a meaningful iterated attack. If we apply it to linear cryptanalysis, this result is thus weaker than Theorem 6. It is however much more general.

2.4 Applications

PEANUT98 is a 9-round Feistel Cipher for message-blocks of size 64 which has been proposed in [13] with a constructed pairwise decorrelation such that $\mathrm{Dec}^2(\mathrm{PEANUT98}) \leq 2^{-76}$ (see Theorem 7). From Theorem 5 we know that no differential distinguisher with a number of chosen plaintext pairs less than 2^{76} will have an advantage greater than 50%. From Theorem 6 we know that no differential distinguisher with a number of known plaintext less than 2^{62} will have an advantage greater than 50%. Now from Theorem 8 we know that no known plaintext iterated attack of order 1 (*e.g.* linear attacks) with a number of known plaintext less than 2^{33} will have an advantage greater than 50%. For linear cryptanalysis, this result is weaker than Theorem 6, but more general.

All these results are applicable to the COCONUT98 Cipher as well since its pairwise decorrelation bias is even smaller (it is actually zero).

3 On Combining Several Attacks

When several (inefficient) attacks hold against a cipher C, it is natural to wonder whether or not we can combine their effort in order to get an efficient attack. This situation is formalized by changing a few things on Fig. 4 and we can rewrite Theorem 8 in this setting. Firstly, the test in each iteration can be changed. Secondly, n must be considered as relatively small, and d as relatively large. The resulting model is illustrated on Fig. 5.

We let now r denotes the number of attacks and $\mathcal{A}_i$ denotes the ith one.

Input: a cipher c, several attacks $\mathcal{A}_1, \ldots, \mathcal{A}_r$, an acceptance set A
1. for i from 1 to r do
   (a) perform the attack $\mathcal{A}_i$
   (b) set T_i to the result of the attack
2. if $(T_1, \ldots, T_r) \in A$ output 1 otherwise output 0

Fig. 5. Combined Attack.

Theorem 9. *Let C be a cipher. Let $\mathcal{A}_1, \ldots, \mathcal{A}_r$ be r attacks on C with advantages $\mathrm{Adv}_1, \ldots, \mathrm{Adv}_r$ respectively. For each i, we let n_i denote the number of queries from $\mathcal{A}_i$ to c and $\mathcal{A}_i^2$ denote the following attack.*

Input: *a cipher c*
1. *perform the attack $\mathcal{A}_i$ and set a to the result*
2. *perform the attack $\mathcal{A}_i$ and set b to the result*
3. *if $a = b = 1$ output 1 otherwise output 0*

We let Adv'_i denote its advantage, and δ_i denote the probability that the two $\mathcal{A}_i$ attack executions query c with one input in common. We assume that we have $n_i \leq M/2$ for any i. For any combined attack (depicted on Fig. 5) with independent attacks, we have

$$\mathrm{Adv}_{\mathrm{Fig.5}}(C) \leq \sum_{i=1}^{r} \left(\mathrm{Adv}_i + 3 \left(2 \left(\delta_i + \frac{5n_i^2}{4M} \right) + 2\mathrm{Adv}_i + \mathrm{Adv}'_i \right)^{\frac{1}{3}} \right).$$

For instance, when the attacks are known plaintext attacks with a plaintext source with uniform distribution, we have $\delta_i \leq \frac{n_i^2}{2M}$.

Proof. As for the proof of Theorem 8, the advantage can be written

$$\mathrm{Adv}_{\mathrm{Fig.5}}(C) = |E(f(Z_1, \ldots, Z_r) - f(Z_1^*, \ldots, Z_r^*))|$$

for a polynomial $f(x_1, \ldots, x_r)$ of partial degrees at most 1 and with values in $[0, 1]$ whenever all entries are in $[0, 1]$. All partial derivatives $f'_i(x_1, \ldots, x_r)$ are in $[-1, 1]$, so we have

$$\mathrm{Adv}_{\mathrm{Fig.5}}(C) \leq \sum_{i=1}^{r} E(|Z_i - Z_i^*|).$$

We have $|E(Z_i - Z_i^*)| = \mathrm{Adv}_i$ and $|E(Z_i^2 - (Z_i^*)^2)| = \mathrm{Adv}'_i$. So, as in the proof of Theorem 8, we obtain

$$\mathrm{Adv}_{\mathrm{Fig.5}}(C) \leq \sum_{i=1}^{r} \left(\mathrm{Adv}_i + 3\left(2V(Z_i^*) + 2\mathrm{Adv}_i + \mathrm{Adv}'_i\right)^{\frac{1}{3}} \right).$$

and finally $V(Z_i^*) \leq \delta_i + \frac{5n_i^2}{4M}$. □

4 Conclusion

We showed how to unify differential and linear distinguishers in a general notion of iterated attack. We then proved that decorrelation enables to quantify the security against any iterated attack. This result happened to be applicable to a real life block cipher. Our result are however not so tight, and it is still an open problem to improve the complexity upper bounds. We encourage researches in this direction.

References

1. Data Encryption Standard. *Federal Information Processing Standard Publication 46*, U. S. National Bureau of Standards, 1977.
2. E. Biham, A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.
3. L. Carter, M. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, vol. 18, pp. 143–154, 1979.
4. H. Feistel. Cryptography and computer privacy. *Scientific American*, vol. 228, pp. 15–23, 1973.
5. H. Gilbert, M. Girault, P. Hoogvorst, F. Noilhan, T. Pornin, G. Poupard, J. Stern, S. Vaudenay. Decorrelated Fast Cipher: an AES candidate. Submitted.
6. M. Luby, C. Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal on Computing*, vol. 17, pp. 373–386, 1988.
7. M. Matsui. Linear cryptanalysis methods for DES cipher. In *Advances in Cryptology EUROCRYPT'93*, Lofthus, Norway, Lectures Notes in Computer Science 765, pp. 386–397, Springer-Verlag, 1994.
8. M. Matsui. The first experimental cryptanalysis of the Data Encryption Standard. In *Advances in Cryptology CRYPTO'94*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 839, pp. 1–11, Springer-Verlag, 1994.
9. K. Nyberg. Perfect nonlinear S-boxes. In *Advances in Cryptology EUROCRYPT'91*, Brighton, United Kingdom, Lectures Notes in Computer Science 547, pp. 378–385, Springer-Verlag, 1991.
10. C. E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, vol. 28, pp. 656–715, 1949.

11. S. Vaudenay. An experiment on DES — Statistical cryptanalysis. In *3rd ACM Conference on Computer and Communications Security*, New Delhi, India, pp. 139–147, ACM Press, 1996.
12. S. Vaudenay. A cheap paradigm for block cipher security strengthening. Technical Report LIENS-97-3, 1997.
13. S. Vaudenay. Provable security for block ciphers by decorrelation. In *STACS 98*, Paris, France, Lectures Notes in Computer Science 1373, pp. 249–275, Springer-Verlag, 1998.
14. S. Vaudenay. Provable security for block ciphers by decorrelation. (Journal Version.) Submitted.
15. S. Vaudenay. The decorrelation technique home-page.
URL:`http://www.dmi.ens.fr/~vaudenay/decorrelation.html`
16. M. N. Wegman, J. L. Carter. New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences*, vol. 22, pp. 265–279, 1981.