Provable Security for Block Ciphers by Decorrelation

Serge Vaudenay

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

Abstract

In this presentation we investigate a new way of protecting block ciphers against classes of attacks (including differential and linear cryptanalysis) which is based on the notion of decorrelation which is fairly connected to Carter-Wegman's notion of universal functions. This defines a simple and friendly combinatorial measurement which enables to quantify the security. We show that we can mix provable protections and heuristic protections. We finally propose two new block ciphers family we call COCONUT and PEANUT, which implement these ideas and achieve quite reasonable performances for real-life applications.

Public research on cryptography has been boosted twenty years ago by the discovery of public-key cryptography. In their seminal papers, Diffie and Hellman [8], Merkle and Hellman [23], Rivest, Shamir and Adleman [30] and others public-key pioneers offered some new directions to researchers. While giving a fairly good promotion to the area of computational number theory and complexity theory (it was not a coincidence that it has been invented shortly after the foundations of NP-completeness theory has been set), it concentrated most of the research energy.

Surprisingly, the area of conventional cryptography was quite older but did not received the same boost. Although the U.S. Government adopted the Data Encryption Standard (DES) in 1977 (the development of which was pushed by banking security needs), the real advances on the research of block ciphers used to be quite rare. Security was almost always based on empiristic approaches: block ciphers were developed in a sufficiently enough complicated way to make their cryptanalysis impossible. Implementation cost however needs the complexity to be not so high, but no one has a sharp idea where to make a trade-off in between. Now, this area of research is moving towards a mature state where we can give results on provable security.

Before the second world war, security of encryption used to be based on the secrecy of the algorithm. Mass telecommunication and computer science networking however pushed the development of public algorithms with secret keys. The most important research result on encryption was found for the application to the telegraph by Shannon in the Bell Laboratories in 1949 [31]. It proved the unconditional security of the Vernam's Cipher which had been published in 1926 [37]. Although quite expensive to implement for networking (because the sender and the receiver need to be synchronized, and it needs quite cumbersome huge keys), this cipher was used in the Red Telephone between Moscow and Washington D.C. during the cold war. Shannon's result also proves that unconditional security cannot be achieved in a better (*i.e.* cheaper) way. For this reason, empiristic security seemed to be the only efficient possibility, and all secret key block ciphers which have been publicly developed were considered to be secure until some researcher published an attack on it. Therefore research mostly grew like a ball game between the designers team and the analysts team and treatment on the general security of block ciphers has hardly been done.

Real advances on the security on block ciphers have been made in the early 90's.

One of the most important result has been obtained by Biham and Shamir in performing a *differential cryptanalysis* on DES [2, 3, 4, 5]. The best version of this attack can recover a secret key with a simple 2^{47}-chosen plaintext attack[1]. Although this attack is heuristic, experiments confirmed the results.

Biham and Shamir's attack was based on statistical cryptanalysis idea which have also been used by Gilbert and Chassé against another cipher [11, 10]. Those ideas inspired Matsui who developed a *linear cryptanalysis* on DES [20, 21]. This heuristic attack, which has been implemented, can recover the key with a 2^{43}-known plaintext attack. Since then, many researchers tried to generalize and improve these attacks (see for instance [18, 17, 15, 34, 16, 24, 35]), but the general ideas was quite the same.

The basic idea of differential cryptanalysis is to use properties like "if x and x' are two plaintext blocks such that $x' = x \oplus a$, then it is likely that $C(x') = C(x) \oplus b$". Then the attack is an iterated two-chosen plaintexts attack which consists in getting the encrypted values of two random plaintexts which verify $x' = x \oplus a$ until a special event like $C(x') = C(x) \oplus b$ occurs. Similarly, the linear cryptanalysis consists in using the probability that $C(x)$ is in a given hyperplane when we know that x is in another given hyperplane, when this probability is far from $1/2$. More precisely, linear cryptanalysis is

[1] So far, the best known attack was an improvement of exhaustive search which requires on average 2^{54} DES computations.

an incremental one-known plaintext attack where we count how many times this event occurs.

Instead of keeping breaking and proposing new encryption functions, some researchers tried to sit down and understand how to protect ciphers against some classes of attacks. Nyberg first formalized the notion of strength against differential cryptanalysis [25], and similarly, Chabaud and Vaudenay formalized the notion of strength against linear cryptanalysis [6]. With this approach we can study how to make internal computation boxes resistant against both attacks. This can be used in a heuristic way by usual active s-boxes counting tricks (see Heys and Tavares [13] for instance). This has also been used to provide provable security against both attacks by Nyberg and Knudsen [26], but in an unsatisfactory way which introduce some algebraic properties which lead to other attacks as shown by Jakobsen and Knudsen [14].

In this presentation, we introduce a new way to protect block ciphers against various kind of attacks. This approach is based on the notion of universal functions introduced by Carter and Wegman [7, 38] for the purpose of authentication. Protecting block ciphers is so cheap that we call NUT (as for "n-Universal Transformation") the set of ciphers which are protected this way. We finally describe two cipher families we call COCONUT (as for "Cipher Organized with Cute Operations and NUT") and PEANUT (as for "Pretty Encryption Algorithm with NUT") and offer two definite examples as a cryptanalysis challenge.

The paper is organized as follows. First we give some definitions an basic constructions for decorrelation. Then we state Shannon's perfect secrecy notion in term of decorrelation. We show how to express security results in the Luby-Rackoff's security model. Then we compute how Feistel Ciphers are decorrelated. We prove how pairwise decorrelation can protect a cipher against Biham-Shamir's differential cryptanalysis and Matsui's cryptanalysis. Finally, we define the COCONUT and PEANUT construction and show how to generate keys for them.

1 Decorrelation

We first give formal definitions of the notion of decorrelation which plays a crucial role in our treatment.

Definition 1. Given a random function F from a given set $\mathcal{A}$ to a given set $\mathcal{B}$ and an integer d, we define the d-wise distribution matrix $[F]^d$ of F as a $\mathcal{A}^d \times \mathcal{B}^d$-matrix where the (x, y)-entry of $[F]^d$ corresponding to the multipoints

$x = (x_1, \ldots, x_d) \in \mathcal{A}^d$ and $y = (y_1, \ldots, y_d) \in \mathcal{B}^d$ 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.

Definition 2. Given two random functions F and G from a given set $\mathcal{A}$ to a given set $\mathcal{B}$, an integer d and a distance D over the vector space $\mathbf{R}^{\mathcal{A}^d \times \mathcal{B}^d}$, we call $D([F]^d, [G]^d)$ the d-wise D-decorrelation between F and G.

A decorrelation of zero means that for any multipoint $x = (x_1, \ldots, x_d)$ the multipoints $(F(x_1), \ldots, F(x_d))$ and $(G(x_1), \ldots, G(x_d))$ have the same distribution, so that F and G are *correlated* in some sense *to the order d*.

It is also important to study the decorrelation of a given random function F to a reference random function. For instance, let say that R is a random function from $\mathcal{A}$ to $\mathcal{B}$ with a uniform distribution. Saying that F and R have a 1-wise decorrelated of zero (or equivalently that F and R are *correlated*) means that for any x_1 the distribution of $F(x_1)$ is uniform. Saying that F is 2-wise correlated to R means that for any $x_1 \neq x_2$ the points $F(x_1)$ and $F(x_2)$ are uniform and independent.

We note that this notion is fairly similar to the notion of universal functions which was introduced by Carter and Wegman [7, 38]. More precisely, we recall that a random function F from $\mathcal{A}$ to $\mathcal{B}$ is ϵ-almost strongly universal$_d$ if for any pairwise different $(x_1, \ldots, x_d)$ and any $(y_1, \ldots, y_d)$ we have

$$\Pr[F(x_i) = y_i; i = 1, \ldots, d] \leq \frac{1}{\#\mathcal{B}^d} + \epsilon.$$

If we define $||A||_\infty = \max_{x,y} |A_{x,y}|$, if F has a d-wise $||.||_\infty$-decorrelated of ϵ to a truly random function, then it is ϵ-almost strongly universal$_d$. The converse is true when $\epsilon \geq \frac{1}{\#\mathcal{B}^d}$. Although the notion is fairly similar, we will use our formalism which is adapted to our context.

For a treatment on block cipher, we consider random permutations C on the block-message space $\mathcal{M}$. (Here the randomness comes from the secret key.) In most of cases, we have $\mathcal{M} = \{0, 1\}^m$. If we define the Perfect Cipher C^* as being a random permutation on $\mathcal{M}$ with a uniform distribution, we are interested in the decorrelation between C and C^*.

For the purpose of our treatment, we define the L_2 norm, the infinity weighted norm N_∞ and the L_∞-associated matrix norm $|||.|||_\infty$ on $\mathbf{R}^{\mathcal{M}^d \times \mathcal{M}^d}$

by:

$$\begin{aligned}
||A||_2 &= \sqrt{\sum_{x,y} (A_{x,y})^2} &(1)\\
N_\infty(A) &= \max_{x,y} \frac{|A_{x,y}|}{\Pr[x \overset{C^*}{\mapsto} y]} &(2)\\
|||A|||_\infty &= \max_x \sum_y |A_{x,y}| &(3)
\end{aligned}$$

where C^* is the Perfect Cipher. For the definition of n_∞ we outline that for any d-wise distribution matrix of a permutation, if $\Pr[x \overset{C^*}{\mapsto} y] = 0$, then $A_{x,y} = 0$ and we can assume $0/0 = 0$.

We recall that the $||.||_2$ and $|||.|||_\infty$ norms are matrix norms, which means that $||A \times B|| \leq ||A||.||B||$. Moreover, N_∞ has a similar property. Namely, for a multipoint $x = (x_1, \ldots, x_d)$, let $\equiv_x$ denotes the equivalence relation among $\{1, \ldots, d\}$ defined by

$$i \equiv_x j \Longleftrightarrow x_i = x_j.$$

We say that a matrix A has a *support of a permutation* distribution matrix if

$$\forall x, y \quad \equiv_x \neq \equiv_y \Longrightarrow A_{x,y} = 0.$$

For the distribution matrices of any cipher, this property holds. Then, for any such matrices A and B, we have $N_\infty(A \times B) \leq N_\infty(A).N_\infty(B)$. Namely, N_∞ is a norm on the sub-algebra of all matrices which have a support of a permutation distribution matrix.

Multiplicativity of the decorrelation is very useful when we consider product ciphers. Concretely, if C_1 and C_2 are two independent ciphers and C^* is the Perfect Cipher, then we have

$$||[C_1 \circ C_2]^d - [C^*]^d|| \leq ||[C_1]^d - [C^*]^d||.||[C_2]^d - [C^*]^d||.$$

(This comes from $[C_i]^d \times [C^*]^d = [C^*]^d$.) This property makes the decorrelation a friendly combinatorial measurement.

2 Basic constructions

Perfect 1-wise decorrelation is easy to achieve when the message-block space $\mathcal{M}$ is given a group structure. For instance we can use $C(x) = x + K$ where K has a uniform distribution on $\mathcal{M}$.

We can construct perfect pairwise decorrelated ciphers on a field structure $\mathcal{M}$ by $C(x) = a.x + b$ where $K = (a, b)$ is uniform in $\mathcal{M}^* \times \mathcal{M}$. This requires to consider the special case $a = 0$ when generating K. On the standard space $\mathcal{M} = \{0, 1\}^m$, it also requires to implement arithmetic on the finite field GF(2^m), which may lead to poor encryption rate on software. As an example we can mention the COCONUT Ciphers (see Section 8).

A similar way to construct perfect 3-wise decorrelated ciphers on a field structure $\mathcal{M}$ is by $C(x) = a + b/(x + c)$ where $K = (a, b, c)$ with $b \neq 0$. (By convention we set $1/0 = 0$.)

Perfect decorrelated ciphers with higher orders require dedicated structure. We can for instance use Dickson's Polynomials.

An alternative way consists of using Feistel Ciphers with decorrelated functions [9]. Given a set $\mathcal{M} = \mathcal{M}_0^2$ where $\mathcal{M}_0$ has a group structure and given r random functions $F_1, \ldots, F_r$ on $\mathcal{M}_0$ we denote $C = \Psi(F_1, \ldots, F_r)$ the cipher defined by $C(x^l, x^r) = (y^l, y^r)$ where we iteratively compute a sequence (x_i^l, x_i^r) such that

$$\begin{aligned} x_0^l &= x^l \quad \text{and} \quad x_0^r = x^r \\ x_i^l &= x_{i-1}^r \quad \text{and} \quad x_i^r = x_{i-1}^l + F_i(x_{i-1}^r) \\ y^l &= x_r^r \quad \text{and} \quad y^r = x_r^l. \end{aligned}$$

(Note that the final exchange between the two halves is cancelled here.) In all the constructions of this paper, $\mathcal{M}_0$ is the group $\mathbf{Z}_2{}^{\frac{m}{2}}$, so the addition is the bitwise exclusive or which will be denoted $\oplus$.

If $\mathcal{M}_0$ has a field structure, we can use perfect d-wise decorrelated F_i functions by $F_i(x) = a_d.x^{d-1} + \ldots + a_2.x + a_1$ where $(a_1, \ldots, a_d)$ is uniformly distributed on $\mathcal{M}_0^d$.

Decorrelation of Feistel Ciphers depends on the decorrelation of all F_i functions. It will be studied in Section 5.

3 Shannon's unconditional security

In this Section, we consider decorrelation of zero. We note that this decorrelation does not depend on the choice of the distance.

Intuitively, if C is 1-wise correlated to C^*, the encryption $C(x_1)$ contains no information on the plaintext-block x_1, so the cipher C is secure if we use it only once (as one-time pad [37]). This corresponds to Shannon's perfect secrecy theory [31]. Similarly, if C is d-wise correlated to C^*, it is unconditionally secure if we use it only d times (on different plaintexts) as the following Theorem shows.

Theorem 3. *Let C be a cipher d-wise correlated to the Perfect Cipher. If an adversary knows $d-1$ pairs $(x_i, C(x_i))$ $(i = 1, \ldots, d-1)$, for any y_d which is different from all $C(x_i)$'s, his knowledge of $C^{-1}(y_d)$ is exactly that it is different from all x_i's. More precisely, for any $x_1, \ldots, x_d$*

$$H(X/C(x_1), \ldots, C(x_d), C(X)) = H(X)$$

where X is a random variable with uniform distribution among all the messages in $\mathcal{M}$ which are different from the x_i's. (Here H denotes Shannon's entropy of random variables.)

We recall that by definition we have $H(X/Y) = H(X,Y) - H(Y)$ and

$$H(X) = -\sum_x \Pr[X = x] \log \Pr[X = x]$$

with the convention that $0 \log 0 = 0$.

Proof. For any x_d which is different from all x_i's, the probability that $C(x_d) = y_d$ knowing all x_i's and $C(x_i)$'s is equal to $\frac{1}{\#\mathcal{M}-r}$ where r is the number of pairwise different x_i's. This is equal to the probability of a uniform X is equal to x_d knowing that it is different from all x_i's. □

4 Security in the Luby-Rackoff model

To illustrate the power of the notion of decorrelation, let us first express the perfect security notion. In the Luby-Rackoff model, an attacker is an infinitely powerful device whose aim is to distinguish a cipher C from the Perfect Cipher C^* by querying an oracle with a limited number n of inputs (see [19]). We assume that we are given an oracle $\mathcal{O}$ which 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 $|p - p^*|$ where p (resp. p^*) is the probability of answering 1 if $\mathcal{O}$ implements C (resp. C^*). We have the following Theorem.

Theorem 4. *If C is a cipher, let d be an integer and ϵ be the d-wise N_∞-decorrelation between C and the Perfect Cipher. For any distinguishing attacker which is limited to d queries, the advantage $|p - p^*|$ is at most ϵ.*

In particular, we have unconditional security for $\epsilon = 0$ and we still have a proven quantified security when ϵ is small.

Proof. Each execution of the attack with an oracle which implements C is characterized by a random tape ω and the successive answers $y_1, \ldots, y_d$ of the queries which we denote $x_1, \ldots, x_d$ respectively. More precisely, x_1 depends on ω, x_2 depends on ω and y_1 and so on. The answer thus depends on $(\omega, y_1, \ldots, y_d)$. Let $\mathcal{E}$ be the set of all $(\omega, y_1, \ldots, y_d)$ such that the output of $\mathcal{A}$ is 1. We have

$$\begin{aligned} p &= \sum_{(\omega,y_1,\ldots,y_d)\in\mathcal{E}} \Pr[\omega] \Pr[C(x_i(\omega, y_1, \ldots, y_{i-1})) = y_i; i = 1, \ldots, d] \\ &\leq (1+\epsilon) \sum_{(\omega,y_1,\ldots,y_d)\in\mathcal{E}} \Pr[\omega] \Pr[C^*(x_i(\omega, y_1, \ldots, y_{i-1})) = y_i; i = 1, \ldots, d] \\ &\leq (1+\epsilon)p^* \end{aligned}$$

so we have $p - p^* \leq \epsilon$ for any attacker. We can apply this result to the attacker which acts as the attack does but produces the opposite output to complete the proof. $\square$

Here is a more precise Theorem in the non adaptive case. (We call an attack "non adaptive" if no x_i queried to the oracle depends on some previous answers y_j.)

Theorem 5. *If C is a cipher, let d be an integer and ϵ be the d-wise $|||.|||_\infty$-decorrelation between C and the Perfect Cipher. The advantage $|p - p^*|$ of the best non-adaptive distinguishing attacker which is limited to d queries is equal to $\epsilon/2$.*

Proof. For those attacks, with the notations of Theorem 4, we have

$$p = \sum_x \Pr[x] \sum_y 1_{(x,y)\in\mathcal{E}} \Pr\left[x \stackrel{C}{\mapsto} y\right]$$

(where 1_P is defined to be 1 if Predicate P is true and 0 otherwise) thus, for the best attack, we have

$$|p - p^*| = \max_{\substack{x\mapsto\Pr[x] \\ \mathcal{E}}} \left| \sum_x \Pr[x] \sum_y 1_{(x,y)\in\mathcal{E}} \left(\Pr\left[x \stackrel{C}{\mapsto} y\right] - \Pr\left[x \stackrel{C^*}{\mapsto} y\right] \right) \right|.$$

We can easily see that this maximum is obtained when $x \mapsto \Pr[x]$ is a Dirac distribution on a multipoint $x = x^0$ and $\mathcal{E}$ includes all y's such that $\Pr\left[x_0 \stackrel{C}{\mapsto} y\right] - \Pr\left[x_0 \stackrel{C^*}{\mapsto} y\right]$ has the same sign, which gives the result. $\square$

5 Decorrelation of Feistel Ciphers

In this Section, we assume that $\mathcal{M} = {\mathcal{M}_0}^2$ where $\mathcal{M}_0$ is a group. Thus we can consider Feistel Ciphers on $\mathcal{M}$.

Theorem 5 can be used in a non-natural way. For instance, let us recall the following Theorem.

Theorem 6 (Luby-Rackoff [19]). *Let F_1, F_2, F_3 be three independent uniform random functions on $\mathcal{M}_0$ and d be an integer. For any distinguishing attacker against $\Psi(F_1, F_2, F_3)$ on $\mathcal{M} = {\mathcal{M}_0}^2$ which is limited to d queries, we have*

$$|p - p^*| \leq \frac{d^2}{\sqrt{\#\mathcal{M}}}.$$

Thus we can say that

$$|||[\Psi(F_1, F_2, F_3)]^d - [C^*]^d|||_\infty \leq 2\frac{d^2}{\sqrt{\#\mathcal{M}}}$$

where C^* is the Perfect Cipher.

For completeness, we also mention some improvements to the previous Theorem due to Patarin [27, 28, 29].

Theorem 7 (Patarin [29]). *Let $F_1, \ldots, F_6$ be six independent uniform random functions on $\mathcal{M}_0$ and d be an integer. For any distinguishing attacker against $\Psi(F_1, \ldots, F_6)$ on $\mathcal{M} = {\mathcal{M}_0}^2$ which is limited to d queries, we have*

$$|p - p^*| \leq \frac{37d^4}{(\#\mathcal{M})^{\frac{3}{2}}} + \frac{6d^2}{\#\mathcal{M}}.$$

So, as Theorem 6 guaranties the security of a three-round Feistel Cipher until $d \sim (\#\mathcal{M})^{\frac{1}{4}}$, this one guaranties the security until $d \sim (\#\mathcal{M})^{\frac{3}{8}}$.

The $|||.|||_\infty$-decorrelation of Feistel Ciphers can be estimated with the following Lemma.

Lemma 8. Let $F_1, \ldots, F_r, R_1, \ldots, R_r$ be $2r$ independent random functions on $\mathcal{M}_0$ such that $|||[F_i]^d - [R_i]^d|||_\infty \leq \epsilon_i$ for $i = 1, \ldots, r$. We have

$$|||[\Psi(F_1, \ldots, F_r)]^d - [\Psi(R_1, \ldots, R_r)]^d|||_\infty \leq (1 + \epsilon_1) \ldots (1 + \epsilon_r) - 1.$$

Proof. Let u^i denotes the input of F_i or R_i in $\Psi(F_1, \ldots, F_r)$ or $\Psi(R_1, \ldots, R_r)$. We thus let (u^0, u^1) denote the input of the ciphers, and (u^{r+1}, u^r) denote the output. Here, all u^i's are multipoints, *i.e.* $u^i = (u^i_1, \ldots, u^i_d)$. We have

$$\begin{aligned}
&\Pr_{F_1,\ldots,F_r}[u^0u^1 \mapsto u^{r+1}u^r] - \Pr_{R_1,\ldots,R_r}[u^0u^1 \mapsto u^{r+1}u^r] \\
&\quad = \sum_{u^2,\ldots,u^{r-1}} \left(\prod_{i=1}^{r} \Pr_{F_i}[u^i \mapsto u^{i-1} \oplus u^{i+1}] - \prod_{i=1}^{r} \Pr_{R_i}[u^i \mapsto u^{i-1} \oplus u^{i+1}] \right) \\
&\quad = \sum_{u^2,\ldots,u^{r-1}} \sum_{\substack{I \subseteq \{1,\ldots,r\} \\ I \neq \emptyset}} \prod_{i \in I} (\Pr_{F_i} - \Pr_{R_i})[u^i \mapsto u^{i-1} \oplus u^{i+1}] \prod_{i \notin I} \Pr_{R_i}[u^i \mapsto u^{i-1} \oplus u^{i+1}]
\end{aligned}$$

hence

$$\begin{aligned}
&\sum_{u^{r+1},u^r} \left| \Pr_{F_1,\ldots,F_r}[u^0u^1 \mapsto u^{r+1}u^r] - \Pr_{R_1,\ldots,R_r}[u^0u^1 \mapsto u^{r+1}u^r] \right| \\
&\quad \leq \sum_{u^2,\ldots,u^{r+1}} \sum_{\substack{I \subseteq \{1,\ldots,r\} \\ I \neq \emptyset}} \prod_{i \in I} |\Pr_{F_i} - \Pr_{R_i}| [u^i \mapsto u^{i-1} \oplus u^{i+1}] \prod_{i \notin I} \Pr_{R_i}[u^i \mapsto u^{i-1} \oplus u^{i+1}] \\
&\quad = \sum_{\substack{I \subseteq \{1,\ldots,r\} \\ I \neq \emptyset}} \prod_{i \in I} \epsilon_i \\
&\quad = (1+\epsilon_1)\ldots(1+\epsilon_r) - 1.
\end{aligned}$$

□

From this Lemma and the previous observation we obtain the following Theorem.

Theorem 9. *Let $F_1, \ldots, F_r, R$ be r independent random functions on $\mathcal{M}_0$ where R has a uniform distribution and such that $|||[F_i]^d - [R]^d|||_\infty \leq \epsilon$ for $i = 1, \ldots, r$. For any $k \geq 3$ we have*

$$|||[\Psi(F_1, \ldots, F_r)]^d - [C^*]^d|||_\infty \leq \left((1+\epsilon)^k - 1 + \frac{2d^2}{\sqrt{\#\mathcal{M}}} \right)^{\lfloor \frac{r}{k} \rfloor}.$$

We can remark that the Lemma remains valid if we replace the group operation used in the Feistel construction by any other (pseudo)group law.

This makes the $|||.|||_\infty$-decorrelation a friendly tool for constructing Feistel Ciphers.

6 Differential cryptanalysis

In this Section we study the security of pairwise decorrelated ciphers against basic differential cryptanalysis. We study criteria which prove that the attack cannot be better than exhaustive attack. In our model, the exhaustive attack for decrypting a given ciphertext y is an attack which exhaustively request for many random $C(x)$'s until we have $C(x) = y$. The complexity of this attack is obviously within the range of the number of possible text blocks.

We assume that $\mathcal{M} = \{0,1\}^m$ is the set of all bitstrings with length m. Let C be a cipher on $\mathcal{M}$ and let C^* be the Perfect Cipher.

Although differential cryptanalysis has been invented in order to recover a whole key by Biham and Shamir (see [4, 5]), we study here the basic underlying notion which makes it work. We call basic differential cryptanalysis the following distinguisher which is characterized by a pair $(a,b) \in \mathcal{M}^2$ with $a \neq 0$.

1. pick a random x with a uniform distribution and query for $C(x)$ and $C(x \oplus a)$
2. if $C(x \oplus a) = C(x) \oplus b$, stop and output 1, otherwise, start again until n trials has been performed
3. stop and output 0

It is well known that differential cryptanalysis depends on the following $\mathrm{DP}^C(a,b)$ (see for instance [25]). We define

$$\mathrm{DP}^C(a,b) = \Pr[C(X \oplus a) = C(X) \oplus b]$$

where X has a uniform distribution. This quantity thus depends on the key. Here we focus on the average value $E\left(\mathrm{DP}^C(a,b)\right)$ over the distribution of the key. We first mention that it has an interesting linear expression with respect to the pairwise distribution matrix of C. Namely, straightforward computation shows that

$$E(\mathrm{DP}^C(a,b)) = 2^{-m} \sum_{\substack{x_1 \neq x_2 \\ y_1 \neq y_2}} 1_{\substack{x_1 \oplus x_2 = a \\ y_1 \oplus y_2 = b}} \Pr\left[\begin{array}{ccc} x_1 & \overset{C}{\mapsto} & y_1 \\ x_2 & \mapsto & y_2 \end{array}\right]. \qquad (4)$$

Lemma 10. For the attack above, we have

$$|p - p^*| \leq n.\max\left(\frac{1}{2^m - 1}, E\left(\mathrm{DP}^C(a,b)\right)\right).$$

So, if $E\left(\mathrm{DP}^C(a,b)\right)$ is within the order of 2^{-m}, then the attack above cannot be better than exhaustive attack.

Proof. It is straightforward to see that the probability, for some fixed key, that the attack accepts C is

$$1 - \left(1 - \mathrm{DP}^C(a,b)\right)^n$$

which is less than $n.\mathrm{DP}^C(a,b)$. Hence we have $p \leq n.E\left(\mathrm{DP}^C(a,b)\right)$. Since from Equation (4) we have $E\left(\mathrm{DP}^{C^*}(a,b)\right) = \frac{1}{2^m-1}$, we obtain the result. $\square$

Theorem 11. *Let $\epsilon = |||[C]^2 - [C^*]^2|||_\infty$ where C^* is the Perfect Cipher. For any basic differential distinguisher between C and C^*, we have $|p - p^*| \leq \frac{n}{2^m-1} + n\epsilon$.*

Proof. Actually we notice that $E\left(\mathrm{DP}^{C^*}(a,b)\right) = \frac{1}{2^m-1}$ and that

$$\left| E\left(\mathrm{DP}^C(a,b)\right) - \frac{1}{2^m - 1} \right| \leq \epsilon$$

from Equation (4). $\square$

So, if ϵ has the order of 2^{-m}, basic differential cryptanalysis does not work against C, but with a complexity in the scale of 2^m.

7 Linear cryptanalysis

Linear cryptanalysis has been invented by Matsui [20, 21] based on the notion of statistical attacks which are due to Gilbert *et al.* [11, 33, 10]. We study here the simpler version of the original attack. With the notations of the previous Section, we similarly call basic linear cryptanalysis the following distinguisher which is characterized by a pair $(a,b) \in \mathcal{M}^2$ with $b \neq 0$.

1. initialize the counter value c to zero
2. pick a random x with a uniform distribution and query for $C(x)$
3. if $x \cdot a = C(x) \cdot b$, increment the counter
4. go to step 2 until n iterations has been performed
5. stop and give an output which only depends on the counter value c

We notice here that the attack depends on the way it accepts or rejects depending on the final counter c value.

Linear cryptanalysis is based on the following quantity as pointed out by Chabaud and Vaudenay [6].

$$\mathrm{LP}^C(a,b) = (2\Pr[X\cdot a = f(X)\cdot b] - 1)^2$$

where $\cdot$ denotes the inner dot product. (We use Matsui's notations taken from [22].) As for differential cryptanalysis, we focus on $E\left(\mathrm{LP}^C(a,b)\right)$, and there is a linear expression of this mean value and the terms of the pairwise distribution matrix $[C]^2$ which comes from straightforward computations :

$$E(\mathrm{LP}^C(a,b)) = 1 - 2^{2-2m}\sum_{\substack{x_1\neq x_2\\ y_1\neq y_2}} 1_{\substack{x_1\cdot a = y_1\cdot b\\ x_2\cdot a\neq y_2\cdot b}} \Pr\left[\begin{array}{ccc} x_1 & \overset{C}{\mapsto} & y_1\\ x_2 & \mapsto & y_2\end{array}\right]. \qquad (5)$$

Lemma 12. For any distinguisher in the above model, we have

$$|p-p^*| \leq 9.3\left(n.\max\left(\frac{1}{2^m-1}, E\left(\mathrm{LP}^C(a,b)\right)\right)\right)^{\frac{1}{3}}.$$

So, if $E\left(\mathrm{LP}^C(a,b)\right)$ is within the order of 2^{-m}, then the attack above cannot be essentially better than exhaustive attack.

Proof. Let N_i be the random variable defined as being 1 or 0 depending on whether or not we have $x\cdot a = C(x)\cdot b$ in the ith iteration. All N_i's are independent and with the same 0-or-1 distribution. Let μ be the probability that $N_i = 1$, for a fixed key K. From the Central Limit Theorem, we can approximate the quantity c/n where c is the counter value to a normal distribution law with mean μ and standard deviation $\sigma = \sqrt{\frac{\mu(1-\mu)}{n}}$. Let A be the set of all accepted c/n quantities. For a fixed key K, the probability that the attack accepts is

$$p^K \approx \int_{t\in A} \frac{e^{-\frac{(t-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}dt.$$

We can compare it to the theoretical expected value p_0 defined by $\mu = \frac{1}{2}$ and $\sigma = \frac{1}{2\sqrt{n}}$. The difference $p^K - p_0$ is maximal when $A = [\tau_1, \tau_2]$ for some values τ_1 and τ_2 which are roots of the Equation

$$\frac{(t-\mu)^2}{\sigma^2} + \log\sigma^2 = 4n\left(t-\frac{1}{2}\right)^2 - \log 4n.$$

Hence, the maximum of the difference $p^K - p_0$ is at most the maximum when $A = [\tau_1, \tau_2]$ over the choice of τ_1 and τ_2, which is the maximum minus the minimum of $p^K - p_0$ when $A =]-\infty, \tau]$. Now we have

$$\int_{-\infty}^{\tau} \frac{e^{-\frac{(t-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dt = \int_{-\infty}^{\frac{\tau-\mu}{\sigma}} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt$$

so we have

$$p^K - p_0 \leq \left(\max_{\tau} - \min_{\tau}\right) \int_{2\sqrt{n}(\tau-\frac{1}{2})}^{\sqrt{n}\frac{\tau-\mu}{\sqrt{\mu(1-\mu)}}} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt.$$

We consider the sum as a function $f(\mu)$ on μ. Since we have $f\left(\frac{1}{2}\right) = 0$, we have $|f(\mu)| \leq B.\left|\mu - \frac{1}{2}\right|$ where B is the maximum of $|f'(x)|$ when x varies from μ to $\frac{1}{2}$. We have

$$f'(x)\sqrt{\frac{2\pi}{n}} = \left(-\frac{1}{\sqrt{x(1-x)}} - \frac{\frac{1}{2}-x}{x(1-x)}\frac{\tau - x}{\sqrt{x(1-x)}}\right) e^{-\frac{n(\tau-x)^2}{2x(1-x)}}$$

so

$$|f'(x)| \leq \sqrt{\frac{n}{2\pi\mu(1-\mu)}} + \frac{\left|\mu - \frac{1}{2}\right|}{\mu(1-\mu)}\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} \leq \sqrt{\frac{n}{2\pi\mu(1-\mu)}} + \frac{\left|\mu - \frac{1}{2}\right|}{\mu(1-\mu)}.$$

Therefore we have

$$\left|p^K - p_0\right| \leq 2\sqrt{\frac{n}{2\pi}}\frac{\left|\mu - \frac{1}{2}\right|}{\sqrt{\mu(1-\mu)}} + 2\frac{\left(\mu - \frac{1}{2}\right)^2}{\mu(1-\mu)}. \tag{6}$$

Let $d = E\left((2\mu - 1)^2\right)$ over the distribution of the key. (We recall that μ depends on the key used in the cipher.) Let $\alpha = \frac{1}{8}\left(d\sqrt{\frac{2\pi}{n}}\right)^{\frac{1}{3}}$. Since $d \leq 1$ and $n \geq 1$, we have $\alpha \leq .17$ so if $\left|\mu - \frac{1}{2}\right| \leq \alpha$ we have

$$\left|p^K - p_0\right| \leq .55(dn)^{\frac{1}{3}}.$$

Now we have $\left|\mu - \frac{1}{2}\right| \geq \alpha$ with a probability less than $\frac{d}{4\alpha^2}$, which is less than $8.68(dn)^{\frac{1}{3}}$, and in this case we have $\left|p^K - p_0\right| \leq 1$. Hence, we have

$$E\left(\left|p^K - p_0\right|\right) \leq 9.3(dn)^{\frac{1}{3}}.$$

We note that $d = E\left(\mathrm{LP}^C(a,b)\right)$. We also have $E\left(\mathrm{LP}^{C^*}(a,b)\right) = \frac{1}{2^m - 1}$ from Equation (5), therefore $|p - p^*|$ is too small if $E\left(\mathrm{LP}^C(a,b)\right)$ has the order of 2^{-m} and $n << 2^m$. □

Theorem 13. *Let $\epsilon = |||[C]^2 - [C^*]^2|||_\infty$ where C^* is the Perfect Cipher. For any basic linear distinguisher between C and C^*, we have $|p - p^*| \leq 9.3\left(\frac{n}{2^m-1} + 4n\epsilon\right)^{\frac{1}{3}}$.*

Proof. Actually we notice that $E\left(\mathrm{LP}^{C^*}(a,b)\right) = \frac{1}{2^m-1}$ and that

$$\left|E\left(\mathrm{LP}^{C}(a,b)\right) - \frac{1}{2^m-1}\right| \leq 4\epsilon$$

from Equation (5). □

So, if ϵ has the order of 2^{-m}, basic linear cryptanalysis does not work against C, but with a complexity in the scale of 2^m.

8 COCONUT: a perfect decorrelation design

In this Section we define the COCONUT Ciphers family which are perfectly decorrelated ciphers to the order two.

The COCONUT Ciphers are characterized by some parameters (m, p). m is the block length, and p is a irreducible polynomial with degree m in GF(2) (which defines a representation of the GF(2^m) Galois Field). A COCONUT Cipher with block length m is simply a product cipher $C_1 \circ C_2 \circ C_3$ where C_1 and C_3 are any (possibly weak) ciphers which can depend from each other, and C_2 is an independent cipher based on a $2m$-bit key which consists of two polynomials A and B with degree at most $m-1$ over GF(2) such that $A \neq 0$. For a given representation of polynomials into m-bit strings, we simply define

$$C(x) = A.x + B \bmod p.$$

Since C_2 performs perfect decorrelation to the order two and since it is independent from C_1 and C_3, any COCONUT Cipher is obviously perfectly decorrelated to the order two. Therefore Theorems 11 and 13 shows that COCONUT resists to the basic differential and linear cryptanalysis.

One can wonder why C_1 and C_3 are for. Actually, C_2 makes some precise attacks provably impractical, but in a way which makes the cipher obviously weak against other attacks. We believe that all real attacks on any real cipher have an intrinsic *order* d, that is they use the d-wise correlation in the encryption of d messages. Attacks with a large d on real ciphers are impractical, because the d-wise decorrelation can hardly be analyzed since it depends on too many factors. Therefore, the COCONUT approach consists in making the cipher provably resistant against attacks with order at most

2 such as differential or linear cryptanalysis, and heuristically secure against attacks with higher order by real life ciphers as C_1 and C_3. The cipher C_2 alone would actually have been very unsecure since two known plaintexts can recover the key which can be used to decrypt any (third) ciphertext.

In the Appendix we propose a concrete example: the COCONUT98 Cipher with parameters $m = 64$ and $p = x^{64} + x^{11} + x^2 + x + 1$.

9 PEANUT: a partial decorrelation design

In this Section we define the PEANUT Ciphers family, which achieves an example of partial decorrelation. This family is based on a combinatorial tool which has been previously used by Halevi and Krawczyk for authentication in [12].

The PEANUT Ciphers are characterized by some parameters (m, r, d, p). They are Feistel Ciphers with block length of m bits (m even), r rounds. The parameter d is the order of partial decorrelation that the cipher performs, and p must be a prime number greater than $2^{\frac{m}{2}}$.

The cipher is defined by a key of $\frac{mrd}{2}$ bits which consists of a sequence of r lists of d $\frac{m}{2}$-bit numbers, one for each round. In each round, the F function has the form

$$F(x) = g(k_1.x^{d-1} + k_2.x^{d-2} + \ldots + k_{d-1}.x + k_d \bmod p \bmod 2^{\frac{m}{2}})$$

where g is any permutation on the set of all $\frac{m}{2}$-bit numbers.

Let us now estimate the $|||.|||_\infty$-decorrelation of the PEANUT ciphers.

Lemma 14. Let $\mathbf{K} = \mathrm{GF}(q)$ be a finite field, let $r : \{0,1\}^{\frac{m}{2}} \to \mathbf{K}$ be an injective mapping, and let $p : \mathbf{K} \to \{0,1\}^{\frac{m}{2}}$ be a surjective mapping. Let F be a random function defined by

$$F(x) = p(r(A_{d-1}).r(x)^{d-1} + \ldots + r(A_0))$$

where the A_i's are independent and uniformly distributed in $\{0,1\}^{\frac{md}{2}}$. If R is an independent uniformly distributed random function, we have

$$|||[F]^d - [R]^d|||_\infty \leq 2\left(\left(\frac{q}{2^{\frac{m}{2}}}\right)^d - 1\right).$$

Proof. Let $x = (x_1, \ldots, x_d)$ be a multipoint in $\{0,1\}^{\frac{m}{2}}$. We want to prove that

$$S = \sum_{y=(y_1,\ldots,y_d)} |[F]^d_{x,y} - [R]^d_{x,y}| \leq 2\left(\left(\frac{q}{2^{\frac{m}{2}}}\right)^d - 1\right).$$

Let c be the number of pairwise different x_i's. For any y such that there exists (i,j) such that $y_i \neq y_j$ and $x_i = x_j$, the contribution to the sum is zero. So we can assume that y is defined over the $2^{\frac{cm}{2}}$ choices of y_i's on positions corresponding to pairwise different x_i's. If we let $x_{d+1}, \ldots, x_{2d-c}$ be new fixed points such that we have exactly d pairwise different x_i's, since the probability that F (resp. R) maps x onto y is equal to the sum over all choices of $y_{d+1}, \ldots, y_{d+c}$ that it maps the extended x onto the extended y, we can assume w.l.o.g. that $c = d$.

For any multipoint y we thus have that $\Pr[x \mapsto y] = j.2^{-\frac{md}{2}}$ where j is an integer. Let N_j be the number of multipoints y which verify this property. We have $\sum_j N_j = 2^{\frac{md}{2}}$ and $\sum_j jN_j = 2^{\frac{md}{2}}$. We have

$$S \leq \sum_j N_j \left| \frac{j-1}{2^{\frac{md}{2}}} \right| = 2N_0.2^{-\frac{md}{2}}.$$

N_0 is the number of unreachable y's, *i.e.* the y's which correspond to a polynomial whose coefficients are not all r-images. This number is thus less than the number of missing polynomials which is $q^d - 2^{\frac{md}{2}}$. $\square$

From Theorem 9 with $k = 3$ we thus obtain the following Theorem.

Theorem 15. *Let C be a cipher in the PEANUT family with parameters (m, r, d, p) and let C^* be the Perfect Cipher. We have*

$$|||[C]^d - [C^*]^d|||_\infty \leq \left(\left(1 + 2\left(p^d 2^{-\frac{md}{2}} - 1\right)\right)^3 - 1 + \frac{2d^2}{2^{\frac{m}{2}}} \right)^{\lfloor \frac{r}{3} \rfloor}.$$

We can thus approximate

$$|||[C]^d - [C^*]^d|||_\infty \approx \left(\frac{6d\left(p - 2^{\frac{m}{2}}\right) + 2d^2}{2^{\frac{m}{2}}} \right)^{\frac{r}{3}}.$$

Example. We can use the parameters $m = 64$, $r = 9$, $d = 2$ and $p = 2^{32}+15$. We obtain that $|||[C]^2 - [C^*]^2|||_\infty \leq 2^{-76}$. Therefore from Theorems 11 and 13 no differential or linear cryptanalysis can efficiently distinguish the cipher from the Perfect Cipher. In the Appendix we propose PEANUT98 which is based on those parameters.

In an earlier version of this work [36], we proposed a similar construction (say PEANUT97) which uses prime numbers smaller than $2^{\frac{m}{2}}$. However the result above does not hold with the $|||.|||_\infty$ norm, but rather with the $||.||_2$ one. The drawback is that this norm has less friendly theorems for constructing Feistel ciphers, and in particular we need more rounds to make the cipher provably secure.

10 Note on the key length

One problem with the COCONUT or PEANUT construction is that it requires a long key (in order to make the internal random functions independent). In real-life examples, we can generate this long key by using a pseudorandom generator fed with a short key, but the results on the security based on decorrelation are no longuer valid. However, provided that the pseudorandom generator produces outputs which are undistinguishable from truly random sequences, we can still prove the security.

Actually, let C be the cipher fed with a key spanned with a short key and let C^* be the cipher fed with a truly random key. We assume that we have a result on the security of C^* based on its decorrelation. If there exists an attack against C which would have contradicted the security if it could be applied against C^*, we can use this attack to distinguish the key spanned by the generator from a random key: the distinguisher just give the output of the attack. Hence if the pseudorandom generator is secure, the security results hold on C as they hold on C^*.

11 Conclusion

Decorrelation modules are cheap and friendly tools which can strengthen the security of block ciphers. Actually, we can quantify their security against a class of cryptanalysis which includes differential and linear cryptanalysis. To illustrate this paradigm, we propose two definite prototype ciphers. (see Appendix). Research on other general cryptanalysis is still an open problem, so we strongly encourage research on analyzing the security of those prototype ciphers.

Acknowledgements

I wish to thank Thomas Pornin and Jacques Stern for valuable help. I also thank the CNRS for having strongly motivated this work.

References

[1] E. Biham. A fast new DES implementation in software. In *Fast Software Encryption*, Haifa, Israel, Lectures Notes in Computer Science 1267, pp. 260–272, Springer-Verlag, 1997.

[2] E. Biham, A. Shamir. Differential cryptanalysis of DES-like cryptosystems. In *Advances in Cryptology CRYPTO'90*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 537, pp. 2–21, Springer-Verlag, 1991.

[3] E. Biham, A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, vol. 4, pp. 3–72, 1991.

[4] E. Biham, A. Shamir. Differential cryptanalysis of the full 16-round DES. In *Advances in Cryptology CRYPTO'92*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 740, pp. 487–496, Springer-Verlag, 1993.

[5] E. Biham, A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.

[6] F. Chabaud, S. Vaudenay. Links between differential and linear cryptanalysis. In *Advances in Cryptology EUROCRYPT'94*, Perugia, Italy, Lectures Notes in Computer Science 950, pp. 356–365, Springer-Verlag, 1995.

[7] L. Carter, M. Wegman. Universal clases of hash functions. *Journal of Computer and System Sciences*, vol. 18, pp. 143–154, 1979.

[8] New directions in cryptography. *IEEE Transactions on Information Theory*, vol. IT-22, pp. 644–654, 1976.

[9] H. Feistel. Cryptography and computer privacy. *Scientific american*, vol. 228, pp. 15–23, 1973.

[10] H. Gilbert. *Cryptanalyse Statistique des Algorithmes de Chiffrement et Sécurité des Schémas d'Authentification*, Thèse de Doctorat de l'Université de Paris 11, 1997.

[11] H. Gilbert, G. Chassé. A statistical attack of the FEAL-8 cryptosystem. In *Advances in Cryptology CRYPTO'90*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 537, pp. 22–33, Springer-Verlag, 1991.

[12] S. Halevi, H. Krawczyk. MMH: software message authentication in the Gbit/second rates. In *Fast Software Encryption*, Haifa, Israel, Lectures Notes in Computer Science 1267, pp. 172–189, Springer-Verlag, 1997.

[13] H. M. Heys, S. E. Tavares. Substitution-Permutation Networks resistant to differential and linear cryptanalysis. *Journal of Cryptology*, vol. 9, pp. 1–19, 1996.

[14] T. Jakobsen, L. R. Knudsen. The interpolation attack on block ciphers. In *Fast Software Encryption*, Haifa, Israel, Lectures Notes in Computer Science 1267, pp. 28–40, Springer-Verlag, 1997.

[15] L. R. Knudsen. *Block Ciphers — Analysis, Design and Applications*, Aarhus University, 1994.

[16] B. R. Kaliski Jr., M. J. B. Robshaw. Linear cryptanalysis using multiple approximations. In *Advances in Cryptology CRYPTO'94*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 839, pp. 26–39, Springer-Verlag, 1994.

[17] X. Lai. *On the Design and Security of Block Ciphers*, ETH Series in Information Processing, vol. 1, Hartung-Gorre Verlag Konstanz, 1992.

[18] X. Lai, J. L. Massey, S. Murphy. Markov ciphers and differential cryptanalysis. In *Advances in Cryptology EUROCRYPT'91*, Brighton, United Kingdom, Lectures Notes in Computer Science 547, pp. 17–38, Springer-Verlag, 1991.

[19] M. Luby, C. Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal on Computing*, vol. 17, pp. 373–386, 1988.

[20] 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.

[21] 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.

[22] M. Matsui. New structure of block ciphers with provable security against differential and linear cryptanalysis. In *Fast Software Encryption*, Cambridge, United Kingdom, Lectures Notes in Computer Science 1039, pp. 205–218, Springer-Verlag, 1996.

[23] R. Merkle, M. Hellman. Hiding information and signatures in trapdoor knapsacks. *IEEE Transactions on Information Theory*, vol. IT-24, pp. 525–530, 1978.

[24] S. Murphy, F. Piper, M. Walker, P. Wild. Likehood estimation for block cipher keys. Unpublished.

[25] 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.

[26] K. Nyberg, L. R. Knudsen. Provable security against a differential cryptanalysis. *Journal of Cryptology*, vol. 8, pp. 27–37, 1995.

[27] J. Patarin. *Etude des Générateurs de Permutations Basés sur le Schéma du D.E.S.*, Thèse de Doctorat de l'Université de Paris 6, 1991.

[28] J. Patarin. In *Advances in Cryptology EUROCRYPT'92*, Balatonfüred, Hungary, Lectures Notes in Computer Science 658, pp. 256–266, Springer-Verlag, 1993.

[29] J. Patarin. About Feistel schemes with six (or more) rounds. To appear in *Fast Software Encryption*, 1998.

[30] R. L. Rivest, A. Shamir, L. M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, vol. 21, pp. 120–126, 1978.

[31] C. E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, vol. 28, pp. 656–715, 1949.

[32] A. Shamir. How to photofinish a cryptosystem? Presented at the Rump Session of Crypto'97.

[33] A. Tardy-Corfdir, H. Gilbert. A known plaintext attack of FEAL-4 and FEAL-6. In *Advances in Cryptology CRYPTO'91*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 576, pp. 172–181, Springer-Verlag, 1992.

[34] S. Vaudenay. *La Sécurité des Primitives Cryptographiques*, Thèse de Doctorat de l'Université de Paris 7, Technical Report LIENS-95-10 of the Laboratoire d'Informatique de l'Ecole Normale Supérieure, 1995.

[35] 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.

[36] S. Vaudenay. A cheap paradigm for block cipher security strengthening. Technical Report LIENS-97-3. Unpublished.

[37] G. S. Vernam. Cipher printing telegraph systems for secret wire and radio telegraphic communications. *Journal of the American Institute of Electrical Engineers*, vol. 45, pp. 109–115, 1926.

[38] 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.

Appendix: COCONUT98 and PEANUT98

We define here two real-life block ciphers in the COCONUT and PEANUT families: COCONUT98 and PEANUT98. Both are on blocks with length 64 bits. They both use a secret key which is defined as a sequence which can be generated by a secure pseudorandom generator.

Without any mention, we identify the bitstrings in $\{0,1\}^\ell$ and the integers in $\{0,\ldots,2^\ell-1\}$ by the natural binary expension mapping

$$b_{\ell-1}\ldots b_1 b_0 \longleftrightarrow 2^{\ell-1}b_{\ell-1}+\ldots+2b_1+b_0.$$

Let first let φ and g be two functions from $\{0,1\}^{32}$ onto itself. We define

$$\varphi(x)=x+256.S(x \bmod 256) \bmod 2^{32}$$

and

$$g(x)=R_L^{11}(\varphi(x))+c \bmod 2^{32}$$

where R_L^{11} is a circular rotation by 11 bits to the left (*i.e.* multiplication by 2^{11} modulo $2^{32}+1$), c is a constant and S is a lookup table of 256 24-bit integers. We use the hexadecimal expansion of the mathematical constant e to define c and S: given that

$$e=\sum_{i=0}^{\infty}\frac{1}{i!}=2.\texttt{b7e15162 8aed2a 6abf71 58809c f4f3c7 62e716}\ldots$$

we define $c=\texttt{b7e15162}$ and S by Tables 1 and 2.

We observe that $x \bmod 256=\varphi(x) \bmod 256$. Thus from $\varphi(x)$ we can recover x by $x=\varphi(x)-256.S(\varphi(x) \bmod 256)$. Since the sum to a constant modulo 2^{32}, R_L and φ are permutations, g is a permutation over the set of 32-bit strings.

The COCONUT98 Cipher

For COCONUT98, we know define

$$f_i(x)=\varphi(g(x\oplus k_i))$$

where $\oplus$ denotes the exclusive or and k_i is a constant defined by the secret key. From the f_i's we can define a Feistel permutation onto the set of 64-bit strings. The COCONUT Cipher consists of two four-round Feistel ciphers $\Psi(f_1,f_2,f_3,f_4)$ and $\Psi(f_5,f_6,f_7,f_8)$ and a decorrelation module in between.

The key space of COCONUT98 is the set of all 256-bit strings K such that if we write the concatenation $K=(K_1,K_2,\ldots,K_8)$ where all K_i's are

	.0	.1	.2	.3	.4	.5	.6	.7
0.	8aed2a	6abf71	58809c	f4f3c7	62e716	0f38b4	da56a7	84d904
1.	bb1185	eb4f7c	7b5757	f59584	90cfd4	7d7c19	bb4215	8d9554
2.	cfbfa1	c877c5	6284da	b79cd4	c2b329	3d20e9	e5eaf0	2ac60a
3.	78e537	d2b95b	b79d8d	caec64	2c1e9f	23b829	b5c278	0bf387
4.	bbca06	0f0ff8	ec6d31	beb5cc	eed7f2	f0bb08	801716	3bc60d
5.	94640d	6ef0d3	d37be6	7008e1	86d1bf	275b9b	241deb	64749a
6.	f10de5	13d3f5	114b8b	5d374d	93cb88	79c7d5	2ffd72	ba0aae
7.	571121	382af3	41afe9	4f77bc	f06c83	b8ff56	75f097	9074ad
8.	5a7db4	61dd8f	3c7554	0d0012	1fd56e	95f8c7	31e9c4	d7221b
9.	c6b400	e024a6	668ccf	2e2de8	6876e4	f5c500	00f0a9	3b3aa7
a.	d1060b	871a28	01f978	376408	2ff592	d9140d	b1e939	9df4b0
b.	c703f5	32ce3a	30cd31	c070eb	36b419	5ff33f	b1c66c	7d70f9
c.	6d8d03	62803b	c248d4	14478c	2afb07	ffe78e	89b9fe	ca7e30
d.	df2be6	4bbaab	008ca8	a06fda	ce9ce7	048984	5a082b	a36d61
e.	558aa1	194177	20b6e1	50ce2b	927d48	d7256e	445e33	3cb757
f.	6b6c79	a58a9a	549b50	c58706	90755c	35e4e3	6b5290	38ca73

Table 1: $S(xy)$ for $y < 8$

	.8	.9	.a	.b	.c	.d	.e	.f
0.	5190cf	ef324e	773892	6cfbe5	f4bf8d	8d8c31	d763da	06c80a
1.	f7b46b	ced55c	4d79fd	5f24d6	613c31	c3839a	2ddf8a	9a276b
2.	cc93ed	874422	a52ecb	238fee	e5ab6a	dd835f	d1a075	3d0a8f
3.	37df8b	b300d0	1334a0	d0bd86	45cbfa	73a616	0ffe39	3c48cb
4.	f45a0e	cb1bcd	289b06	cbbfea	21ad08	e1847f	3f7378	d56ced
5.	47dfdf	b96632	c3eb06	1b6472	bbf84c	26144e	49c2d0	4c324e
6.	7277da	7ba1b4	af1488	d8e836	af1486	5e6c37	ab6876	fe690b
7.	9a787b	c5b9bd	4b0c59	37d3ed	e4c3a7	939621	5edab1	f57d0b
8.	bed0c6	2bb5a8	7804b6	79a0ca	a41d80	2a4604	c311b7	1de3e5
9.	e6342b	302a0a	47373b	25f73e	3b26d5	69fe22	91ad36	d6a147
a.	e14ca8	e88ee9	110b2b	d4fa98	eed150	ca6dd8	932245	ef7592
b.	391810	7ce205	1fed33	f6d1de	9491c7	dea6a5	a442e1	54c8bb
c.	60c08f	0d61f8	e36801	df66d1	d8f939	2e52ca	ef0653	199479
d.	1e99f2	fbe724	246d18	b54e33	5cac0d	d1ab9d	fd7988	a4b0c4
e.	2b3bd0	0fb274	604318	9cac11	6cedc7	e771ae	0358ff	752a3a
f.	3fd1aa	a8dab4	0133d8	0320e0	790968	c76546	b993f6	c8ff3b

Table 2: $S(xy)$ for $y \geq 8$

32-bit strings, not all the 64 bits of (K_7, K_8) are set to zero. We define the k_i's by

i	1	2	3	4
k_i	K_1	$K_1 \oplus K_3$	$K_1 \oplus K_3 \oplus K_4$	$K_1 \oplus K_4$

i	5	6	7	8
k_i	K_2	$K_2 \oplus K_3$	$K_2 \oplus K_3 \oplus K_4$	$K_2 \oplus K_4$

which can easily be performed in hardware: in the first (resp. second) Feistel cipher, we first load K_1 (resp. K_2) in a round-key register and at each round xor it with K_3 or K_4 alternately. The decorrelation module is a function M defined by $(K_5, \ldots, K_8)$ by

$$M(xy) = (xy \oplus K_5K_6) \times K_7K_8$$

where the product is performed in GF(2^{64}). The Galois Field representation is chosen so that a 64-bit string $b_{63} \ldots b_1b_0$ represents the polynomial

$$b_{63}.x^{63} + \ldots + b_1.x + b_0$$

modulo $x^{64} + x^{11} + x^2 + x + 1$ modulo 2.

Finally, the COCONUT98 Cipher is defined by

$$C_K(xy) = \Psi(f_5, f_6, f_7, f_8)\,(M\,(\Psi(f_1, f_2, f_3, f_4)(xy)))$$

which is illustrated on Figure 1. We assume that the secret key is uniformly distributed among the bitstrings such that $K_7K_8 \neq 0$.

We note that the decryption can be performed by using a key K^{-1} defined by

$$K^{-1} = (K_2 \oplus K_4, K_1 \oplus K_4, K_3, K_4, K'_5, K'_6, K'_7, K'_8)$$

with

$$\begin{aligned} K'_5K'_6 &= K_5K_6 \times K_7K_8 \\ K'_7K'_8 &= 1/K_7K_8 \end{aligned}$$

in GF(2^{64}).

The PEANUT98 Cipher

PEANUT98 is in the PEANUT family with parameters $(64, 9, 2, 2^{32}+15)$. It is thus a 9-rounds Feistel Cipher. The secret key is a uniformly distributed 576-bit string which is split into 32-bit strings $K = (K_1, \ldots, K_{18})$. We define the round function in the i-th round as

$$f'_i(x) = g(x.K_{2i-1} + K_{2i} \bmod 2^{32} + 15 \bmod 2^{32}).$$

Decryption consists of inverting the order of the 64-bit blocks in the key string, *i.e.* flipping K_1 and K_{17}, K_2 and K_{18}, K_3 and K_{15}, *etc.*

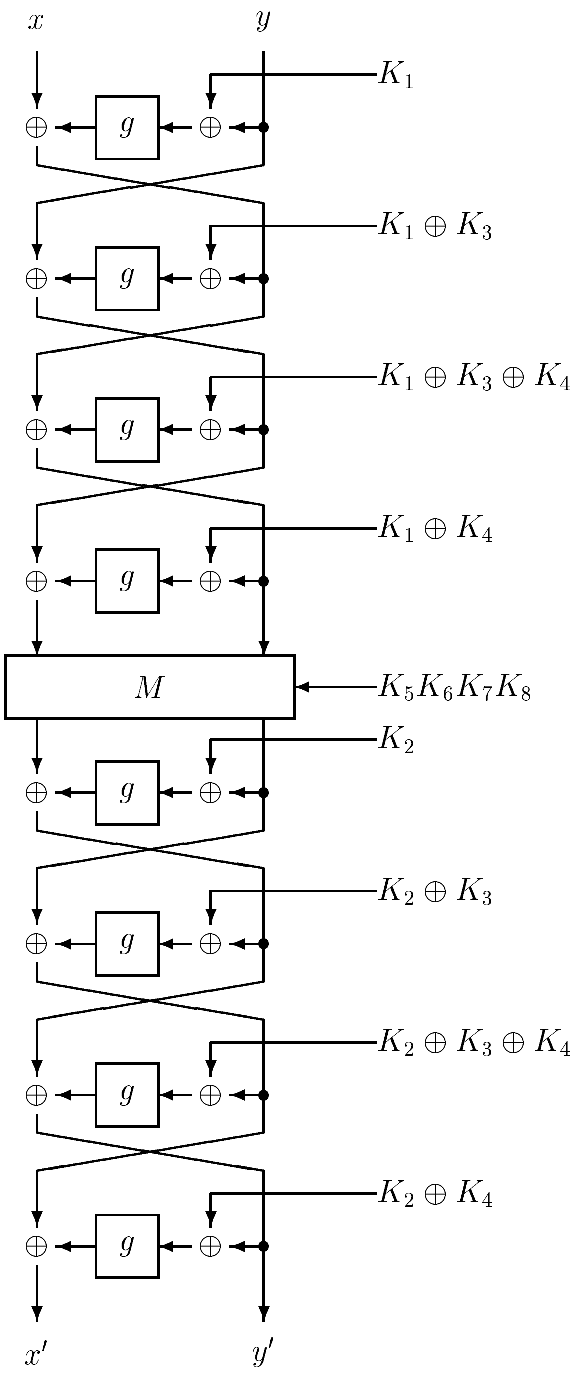

Figure 1: The COCONUT Cipher

Note. We remark that although the PEANUT construction leads to provably secure ciphers against previously mentioned attacks in average among all the keys, it may be possible that some small set of keys are weak. It may be unsecure that there exists some set of keys with density $1/p$ such that there are some simple attacks on when p is not too large. For instance, the PEANUT98 Cipher may accept all the keys such that for some i we have $K_{2i-1} = 0$ as weak keys. Although we did not try to describe an attack in full details, we suggest that those keys shall not be used.

Test Values

As an example, we use the secret key

$$\begin{aligned} K &= \texttt{7c44a4ad56f6bb77220e96e8401694e1} \\ & \texttt{6c469dbc516decc517929e9b226ddd64} \end{aligned}$$

to encrypt the plaintext `6d8779e078ac5f02` with COCONUT98. Here are the intermediate values of the left and right halves in the encryption:

round#1 :	78ac5f02	62b29ee4
round#2 :	62b29ee4	f257ec80
round#3 :	f257ec80	2c554933
round#4 :	6d40ec0e	2c554933
decorrelation :	9c7e2827	751e12b5
round#5 :	751e12b5	ed8378b4
round#6 :	ed8378b4	4f8ba7ff
round#7 :	4f8ba7ff	037910f8
round#8 :	3b2ae895	037910f8

so the ciphertext is `3b2ae895037910f8`. For instance, in the first round we have to compute

$$\begin{aligned} g(\texttt{78ac5f02} \oplus \texttt{7c44a4ad}) &= g(\texttt{04e8fbaf}) \\ &= \varphi(R_L^{11}(\varphi(\texttt{04e8fbaf})) + c) \\ &= \varphi(R_L^{11}(\texttt{f45e8daf}) + c) \\ &= \varphi(\texttt{f46d7fa2} + c) \\ &= \varphi(\texttt{f46d7fa2} + \texttt{b7e15162}) \\ &= \varphi(\texttt{ac4ed104}) \\ &= \texttt{0f35e704} \end{aligned}$$

which is xored onto `6d8779e0` to produce `62b29ee4`.

As an example for PEANUT98, we use the secret key

$$\begin{aligned}
k_1 &= \texttt{2115e265} \\
k_2 &= \texttt{9225cb79} \\
k_3 &= \texttt{cfa1c6fc} \\
k_4 &= \texttt{bd67eef1} \\
k_5 &= \texttt{58cb0b8f} \\
k_6 &= \texttt{fbf151b1} \\
k_7 &= \texttt{423c41e6} \\
k_8 &= \texttt{ec11b5d9} \\
k_9 &= \texttt{b9002c83} \\
k_{10} &= \texttt{406c0b46} \\
k_{11} &= \texttt{ba977fbd} \\
k_{12} &= \texttt{91c0adf4} \\
k_{13} &= \texttt{5b716ec6} \\
k_{14} &= \texttt{1533a950} \\
k_{15} &= \texttt{080b807e} \\
k_{16} &= \texttt{a1a305e3} \\
k_{17} &= \texttt{2a0f096e} \\
k_{18} &= \texttt{4b027140}
\end{aligned}$$

The encryption of `0123456789abcdef` is `07f141edac6485df`. In the first round, we have to compute

$$\begin{aligned}
f'_1(\texttt{89abcdef}) &= g(\texttt{89abcdef}.\texttt{2115e265} + \texttt{9225cb79} \bmod 2^{32} + 15) \\
&= g(\texttt{037724cf7}) \\
&= R_L^{11}(\texttt{703cbff7}) + c \bmod 2^{32} \\
&= \texttt{e5ffbb81} + c \bmod 2^{32} \\
&= \texttt{9de10ce3}
\end{aligned}$$

which is xored onto `01234567` to produce `9cc24984`.

Implementation

Because of the finite field multiplication, the COCONUT98 Cipher is adapted to hardware implementations. Software implementations may require dedicated tricks such as partial table look-up for the multiplications. Biham's bit-slice parallel implementation technics may be a good way to solve the problem too (see [1]).

The design of PEANUT98 is well adapted to software implementation on microprocessors enable to perform 32-bit integer multiplications. A software implementation on a Pentium gave a short (non-optimized) code with only 396 clock cycles per block encryption. This leads to a 20.5Mbps encryption rate at 133MHz.

The integer multiplication makes Biham's bit-slice parallel programming method impractical (see [1]). This has however a good consequence on its security because it makes Shamir's photo-finishing attack impossible too [32].