On the Security of CS-Cipher

Serge Vaudenay

Ecole Normale Supérieure — CNRS
e-mail: Serge.Vaudenay@ens.fr

Abstract. CS-Cipher is a block cipher which has been proposed at FSE 1998. It is a Markov cipher in which diffusion is performed by multipermutations. In this paper we first provide a formal treatment for differential, linear and truncated differential cryptanalysis, and we apply it to CS-Cipher in order to prove that there exists no good characteristic for these attacks. This holds under the approximation that all round keys of CS-Cipher are uniformly distributed and independent. For this we introduce some new technique for counting active Sboxes in computational networks by the Floyd-Warshall algorithm.

Since the beginning of modern public research in symmetric encryption, block ciphers are designed with fixed computational networks: we draw a network and put some computation boxes on. The Feistel scheme [13] is a popular design which enables to make an invertible function with a random function. Its main advantage is that decryption and encryption are fairly similar because we only have to reverse the order of operations.

Another popular (and more intuitive) design consists of having a cascade of computational layers, some of which implement parallel invertible transformations. (People inappropriately call it the "SPN structure" as for Substitution Permutation Network, as opposed to Feistel schemes. Referring to Adams' Thesis [3], several Feistel schemes are also SPN ones.) For this we need two different implementations for encryption and decryption. Several such designs have been proposed to the Advanced Encryption Standard process: Serpent, Safer+, Rijndael and Crypton (see [2]). In this paper we focus on CS-Cipher [32] in order to investigate its security.[1]

The main general known attacks are Biham and Shamir's differential cryptanalysis [8] and Matsui's cryptanalysis [23, 24]. Over their variants, Knudsen's truncated differentials [18, 19] have been shown to be powerful against Massey's Safer block cipher [22], so we investigate it as well. In this paper we consider these attacks and we (heuristically) show that CS-Cipher is resistant against it. For this we use the well known active Sboxes counting arguments techniques.

Here we first recall what can be formally proven under the intuitive approximation that all round keys are uniformly distributed and independent for

[1] While this paper was presented, the owner of the CS-Cipher algorithm announced a "Challenge CS-Cipher": a 10000 euros award will be given to the first person who will decrypt a message encrypted with a key which has been purposely limited to 56 bits. This is basically an exhaustive search race. See http://www.cie-signaux.fr/.

differential and linear cryptanalysis. We contribute to a new similar analysis of truncated differential cryptanalysis. Then we apply these techniques to CS-Cipher. In particular we show how to count the minimal number of active Sboxes in a computational network with multipermutations by using some easy graph algorithms.

1 Previous Work

Public research on cryptography arose in the late 70s. On block ciphers, research has been paradoxically motivated by the Data Encryption Standard [1] controversial: the fact that the design rationales of DES was kept secret by the US government.

Originally the research community was focusing on the fascinating nonlinear properties of the DES Sboxes (and on the existence of a mythical hidden trapdoor). Nonlinear criterion has been investigated, and the possibility on how to achieve it (see Adams and Tavares [3, 4], Nyberg [27]).

Differential cryptanalysis has been invented by Biham and Shamir in the 90s [7], and a connection between the security against it and the nonlinearity of the Sboxes has been found (see Nyberg [28]). Later on the same link arose with Matsui's linear cryptanalysis [23, 24] (see Nyberg [29] and Chabaud-Vaudenay [9]).

Since then an important effort has been done in order to study the security of block ciphers against differential and linear cryptanalysis.

Lai and Massey first invented the notion of "Markov cipher" which enables to make a formal treatment on the security against differential cryptanalysis (see [21, 20]). This enables to formally prove some heuristic approximations used in this attack.

It was well known that the resistance against differential cryptanalysis depends on the minimal number of "active Sboxes" in a characteristic. (The bulk of Biham and Shamir's attack against DES is to find a characteristic with a number of active Sboxes as small as possible.)

In Heys and Tavares [16, 17] is defined the notion of "diffusion order" which enables to get lower bound on the number of active Sboxes in a substitution-permutation network. Similarly, Daemen [11] talks about "branch number". The question was also addressed by Youssef, Mister and Tavares [40]. These notions can be used together with the diffusion properties of the network. Namely, when we use "multipermutations" (see [33, 34]) we can compute these numbers. In particular, inspired by the notion of multipermutation, Daemen, Knudsen and Rijmen use an MDS code in the Square cipher [12] which has been used for two AES candidates: Rijndael and Crypton [2].

An alternate way to prove the security against differential and linear cryptanalysis is to use the Theorem of Nyberg and Knudsen [30, 31] (or a variant), which has been done by Matsui in the Misty cipher [25, 26]. We can also use the "decorrelation theory" [35–39] which has been used in order to create the Peanut and Coconut cipher families (see [15, 37]) and the DFC cipher [15] which is an AES candidate [2] (see [38, 5]). Both approachs provide provable security against differential and linear cryptanalysis (and not only a heuristic security).

2 Formal Treatment on Markov Ciphers

In this section we consider an r-round cipher

$$\mathrm{Enc}(x) = (\rho_r \circ \ldots \circ \rho_1)(x)$$

in which each ρ_i round uses a subkey k_i. We assume that this is a Markov cipher with respect to the XOR addition law, which means, following Lai [20] that for any round i and any x, a and b, we have

$$\Pr_{k_i}[\rho_i(x \oplus a) \oplus \rho_i(x) = b] = \Pr_{k_i, X}[\rho_i(X \oplus a) \oplus \rho_i(X) = b]$$

where X is uniformly distributed. (Here $\oplus$ denotes the bitwise XOR operation.)

2.1 Preliminaries

For any p-to-q-bit function f, any p-bit a and any q-bit b, let us denote

$$\mathrm{DP}^f(a, b) = \Pr_X[f(X \oplus a) \oplus f(X) = b] \tag{1}$$

$$\mathrm{LP}^f(a, b) = \left(2 \Pr_X[a \cdot X = b \cdot f(X)] - 1\right)^2. \tag{2}$$

(Here $a \cdot x$ denotes the dot product of a and x: the sum modulo 2 of all $a_i x_i$.) It is well known that we have

$$\mathrm{LP}^f(a, b) = 2^{-p} \sum_{x,y} (-1)^{(a \cdot x)+(b \cdot y)} \mathrm{DP}^f(x, y) \tag{3}$$

$$\mathrm{DP}^f(a, b) = 2^{-q} \sum_{x,y} (-1)^{(a \cdot x)+(b \cdot y)} \mathrm{LP}^f(x, y). \tag{4}$$

For any random function F (or equivalently any function f which depends on a random parameter K) we consider the expected values over the distribution of F

$$\mathrm{EDP}^F(a, b) = E_F(\mathrm{DP}^F(a, b)) \tag{5}$$

$$\mathrm{ELP}^F(a, b) = E_F(\mathrm{LP}^F(a, b)). \tag{6}$$

Obviously, Equations similar to (3) and (4) hold for EDP and ELP.

2.2 Differential Cryptanalysis of Markov Ciphers

Biham and Shamir's original differential cryptanalysis (see [8]) is defined by a characteristic

$$\Omega = (\omega_0, \ldots, \omega_r) \tag{7}$$

and focus on the probabilistic event

$$E_\Omega : \{M_i \oplus M_i' = \omega_i ; i = 0, \ldots, r / M_0 \oplus M_0' = \omega_0\}$$

where M_0 and M_0' are two different plaintexts and M_i and M_i' are the image of M_0 and M_0' respectively by

$$\rho_i \circ \ldots \circ \rho_1.$$

We let $\Delta M_i = M_i \oplus M_i'$.

Differential cryptanalysis uses the E_Ω event by looking at random pairs (M_0, M_0') such that $\Delta M_0 = \omega_0$ until E_Ω occurs. Thus we try n random pairs, the success rate is at most the probability that one out of the n pairs makes the E_Ω event occur. This probability is

$$1 - \left(1 - \Pr_{M_0, M_0'}[E_\Omega]\right)^n. \tag{8}$$

Thus the probability of success is less than $n \Pr[E_\Omega]$. The average probability of success (over the distribution of the key) is less than

$$n E_k \left(\Pr_{M_0, M_0'}[E_\Omega]\right).$$

Thus we need a number of trials of $p / E(\Pr[E_\Omega])$ in order to achieve an average probability of success at least p.

We define the following formal product

$$\mathrm{DP}(\Omega) = \prod_{i=1}^{r} \mathrm{DP}^{\rho_i}(\omega_{i-1}, \omega_i)$$

(which depends on the key) and

$$\mathrm{EDP}(\Omega) = \prod_{i=1}^{r} \mathrm{EDP}^{\rho_i}(\omega_{i-1}, \omega_i)$$

(which does not). We have the following result which is fairly similar to the treatment of Lai, Massey and Murphy [21].[2]

[2] The difference is that these authors assume the "principle of stochastic equivalence" which enables to remove the expectation over the distribution of the key. The proof is exactly the same.

Lemma 1. *If* Enc *is a Markov cipher and if the round keys $k_1, \ldots, k_r$ are uniformly distributed and independent, we have*

$$E_k\left(\Pr_{M_0,M_0'}[E_\Omega]\right) = \mathrm{EDP}(\Omega).$$

Proof. Since Enc is a Markov cipher and that the keys are independent, from Lai [20] (see also [21]) we know that $\Delta M_0, \ldots, \Delta M_r$ is a Markov chain. Thus we have

$$\begin{aligned}
E_k\left(\Pr_{M_0,M_0'}[E_\Omega]\right) &= \prod_{i=1}^{r} \Pr_{M_0,M_0',k}[\Delta M_i = \omega_i / \Delta M_{i-1} = \omega_{i-1}, \ldots, \Delta M_0 = \omega_0] \\
&= \prod_{i=1}^{r} \Pr_{M_0,M_0',k}[\Delta M_i = \omega_i / \Delta M_{i-1} = \omega_{i-1}] \\
&= \prod_{i=1}^{r} \mathrm{EDP}^{\rho_i}(\omega_{i-1}, \omega_i) \\
&= \mathrm{EDP}(\Omega).
\end{aligned}$$

□

We thus have the following theorem.

Theorem 2. *Given a Markov cipher* $\mathrm{Enc} = \rho_r \circ \ldots \circ \rho_1$ *for the XOR addition law which uses r independent round keys and any differential characteristic $\Omega = (\omega_0, \ldots, \omega_r)$, in order to achieve an average probability of success greater than p for a differential cryptanalysis we need a minimum number of trials of at least $p\,(\mathrm{EDP}(\Omega))^{-1}$. This holds in the model were the probability of success of the differential cryptanalysis for a fixed key is given by Equation (8).*

We emphasis that this is a real formal theorem which does not relies on unproven assumptions.

2.3 Truncated Differential Cryptanalysis

In Biham and Shamir's original differential cryptanalysis, we have a given input difference ω_0 and we expect a given output difference ω_r when the computation follows a path of differences $\omega_1, \ldots, \omega_{r-1}$. It is sometimes useful to consider a multi-path with same ω_0 and ω_r. Actually, we have

$$\mathrm{DP}^{\mathrm{Enc}}(\omega_0, \omega_r) = \sum_{\omega_1,\ldots,\omega_{r-1}} \Pr_{M_0,M_0'}[E_{\omega_0,\ldots,\omega_r}]. \tag{9}$$

Thus, from Lemma 1, we obtain that

$$\mathrm{EDP}^{\mathrm{Enc}}(\omega_0, \omega_r) = \sum_{\omega_1,\ldots,\omega_{r-1}} \prod_{i=1}^{r} \mathrm{EDP}^{\rho_i}(\omega_{i-1}, \omega_i). \tag{10}$$

In the original differential cryptanalysis we only consider the overwhelming term of this sum.

An alternative way is to consider a sub-sum of characteristics which correspond to the same pattern. For instance, Knudsen's truncated differentials [18] corresponds to the sum of all characteristics which predict some part of the differences, *i.e.* for which

$$\forall j \notin I_i \quad (\omega_i)_j = d_{i,j}.$$

In most of block ciphers, what makes the probability of a characteristic small is the differences $d_{i,j}$ of zero. We thus focus on the propagation of zero differences. (Actually, the attack on Safer by Knudsen and Berson [19] uses truncated differentials with zeroes.) Let us denote

$$\mathrm{Supp}(\omega_i) = \{j; (\omega_i)_j \neq 0\}.$$

If we focus on characteristics in which

$$\forall i \quad \mathrm{Supp}(\omega_i) = I_i \text{ and } \omega_i \in A_i$$

we can get a maximal probability with largest A_i sets. The multi-path sum is thus defined by $\Omega = (I_0, \ldots, I_r)$ and we consider the event

$$E_\Omega = \{\mathrm{Supp}(\Delta M_i) = I_i; i = 1, \ldots, r / \mathrm{Supp}(\Delta M_0) = I_0\}$$

in which $\Omega = (I_0, \ldots, I_r)$. We call Ω a "support characteristic".

Theorem 3. *Given a Markov cipher* $\mathrm{Enc} = \rho_r \circ \ldots \circ \rho_1$ *for the XOR addition law which uses* r *independent round keys we consider a truncated differential cryptanalysis. We heuristically assume that there is an overwhelming support characteristic* $\Omega = (I_0, \ldots, I_r)$ *which is such that the probability of success of the differential cryptanalysis for a fixed key is given by Equation (8). The complexity of the attacks must be greater than*

$$p \left(\prod_{i=1}^{r} \Pr_{M_{i-1}, M'_{i-1}, k_i} [\mathrm{Supp}(\Delta M_i) = I_i / \mathrm{Supp}(\Delta M_{i-1}) = I_{i-1}] \right)^{-1}$$

in order to get an average probability of success greater than p*, with the notations of Section 2.2.*

2.4 Linear Cryptanalysis

Linear cryptanalysis is fairly similar to differential cryptanalysis. Here we consider a characteristic Ω associated with a set of linear approximations

$$(\omega_i \cdot M_i) \oplus (\omega_{i+1} \cdot M_{i+1}) \approx \alpha_i \cdot k_i.$$

The characteristic is not associated with a particular event, but corresponds to an (assumed) overwhelming term in the multi-path sum. We define the following formal product

$$\mathrm{LP}(\Omega) = \prod_{i=1}^{r} \mathrm{LP}^{\rho_i}(\omega_{i-1}, \omega_i)$$

(which depends on the key) and

$$\mathrm{ELP}(\Omega) = \prod_{i=1}^{r} \mathrm{ELP}^{\rho_i}(\omega_{i-1}, \omega_i).$$

We have the following result.

Lemma 4. *If* Enc *is a Markov cipher and if the round keys* $k_1, \ldots, k_r$ *are independent, we have*

$$\mathrm{ELP}^{\mathrm{Enc}}(\omega_0, \omega_r) = \sum_{\omega_1, \ldots, \omega_{r-1}} \mathrm{ELP}(\omega_0, \ldots, \omega_r).$$

Proof. First, by Equations (3) and (10) we have

$$\mathrm{ELP}^{\mathrm{Enc}}(\omega_0, \omega_r) = 2^{-\ell} \sum_{u_0, \ldots, u_r} (-1)^{(\omega_0 \cdot u_0) + (\omega_r \cdot u_r)} \prod_{i=1}^{r} \mathrm{EDP}^{\rho_i}(u_{i-1}, u_i)$$

where ℓ is the bit-length of the plaintext. If we now use Equation (4), after a few formal computation steps we obtain the result. □

Matsui's original linear cryptanalysis assumes that one characteristic in the sum is overwhelming, and the attack has a heuristic complexity equal to the inverse of $\mathrm{ELP}^{\mathrm{Enc}}(\omega_0, \omega_r)$. We can thus get a (heuristic) complexity lower bound by upper bounding $\mathrm{ELP}(\Omega)$.

3 On the Security of CS-Cipher

3.1 Presentation of CS-Cipher

In this paper we use non standard notations for CS-Cipher which are better adapted for our treatment. We recall that the secret key is first transformed into a 64-bit subkey sequence $k^0, \ldots, k^8$. We also use two 64-bit constants c and c'. We let $k_0, \ldots, k_{24}$ denote the sequence

$$k^0, c, c', k^1, c, c', \ldots, k^7, c, c', k^8.$$

We thus consider a modified CS-Cipher which is denoted CSC* which is defined by a 1600-bit random key $k = (k_0, \ldots, k_{24})$ with a uniform distribution.

Let us denote

$$s_i(x) = x \oplus k_i$$

which is thus used to "randomize" the message block with a subkey.

We split the standard mixing function M into one linear transformation μ and two involutions P. The linear mapping μ takes two 8-bit inputs and produces two 8-bit outputs by

$$\mu(a,b) = (\varphi(a) \oplus b, R_l(a) \oplus b).$$

Here R_l is a circular rotation by one position to the left, and φ is the standard CS-Cipher operation defined by

$$\varphi(x) = (R_l(x) \wedge \mathtt{55}) \oplus x$$

where $\wedge$ denotes the bitwise AND operation and `55` is the 8-bit hexadecimal constant `01010101`. For convenience we let μ^4 the linear mapping which takes eight 8-bit inputs and produces eight 8-bit outputs by four parallel μ operations

$$\mu^4(x_1,\ldots,x_8) = (\mu(x_1,x_2),\ldots,\mu(x_7,x_8)).$$

We let P denotes the standard CS-Cipher involution defined by a table look-up, and P^8 the application of eight parallel P computations:

$$P^8(x_1,\ldots,x_8) = (P(x_1),\ldots,P(x_8)).$$

We know let L_π denote the following permutation

$$L_\pi(x_1,\ldots,x_8) = (x_1,x_3,x_5,x_7,x_2,x_4,x_6,x_8).$$

We let

$$\rho_i = L_\pi \circ P^8 \circ \mu^4 \circ s_{i-1}$$

for $i = 1,\ldots,23$ and

$$\rho_{24} = s_{24} \circ L_\pi \circ P^8 \circ \mu^4 \circ s_{23}$$

One CS-Cipher block encryption is defined by

$$\mathrm{Enc} = \rho_r \circ \ldots \circ \rho_1.$$

This way we can consider CSC* of being a 24-round cipher in which each round consists of one subkey offset, the μ^4 linear mixing function, the P^8 confusion boxes and the L_π permutation. Due to the s_i structure, it is obvious that CSC* is a Markov cipher.

3.2 Differential Cryptanalysis

We consider a differential characteristic $\Omega = (\omega_0,\ldots,\omega_{24})$ and we aim to upper bound $\mathrm{EDP}(\Omega)$.

We compute each term of the product in $\mathrm{EDP}(\Omega)$. Since s_i, L_π and μ^4 are linear, we let

$$\delta_i = \mu^4(\omega_{i-1}) \tag{11}$$

and

$$\delta'_i = (L_\pi)^{-1}(\omega_i). \tag{12}$$

δ_i and δ'_i are the input and output differences of the ith P^8 layer respectively. From the linearity of μ^4 and L_π and the parallelism of P^8 we obtain

$$\mathrm{EDP}(\Omega) = \prod_{i=1}^{24}\prod_{j=1}^{8} \mathrm{DP}^P((\delta_i)_j, (\delta'_i)_j). \tag{13}$$

Since P is a permutation, the following assumption is necessary for having $\mathrm{EDP}(\Omega) \neq 0$

$$\forall i \;\; \mathrm{Supp}(\delta_i) = \mathrm{Supp}(\delta'_i). \tag{14}$$

We say that a P-box corresponding to indices i, j is "active" if $(\delta_i)_j \neq 0$. We use the following definition.

Definition 5. *For any differential characteristic $\Omega = (\omega_0, \ldots, \omega_{24})$, we define the δ_is and δ'_is by Equations (11) and (12) respectively. We say that Ω is "consistent" if the property of Equation (14) holds. Let $\#\Omega$ denotes the number of indices i, j such that $(\delta_i)_j \neq 0$.*

We thus have the following result.

Lemma 6. *For any non-zero differential characteristic Ω we have*

$$\mathrm{EDP}(\Omega) \leq (\mathrm{DP}^P_{\max})^{\#\Omega}$$

where $\mathrm{DP}^P_{\max} = 2^{-4}$ and $\mathrm{EDP}(\Omega) = 0$ if Ω is not consistent.

Proof. We start from Equation (13). If Ω is not consistent, we obviously have $\mathrm{EDP}(\Omega) = 0$ since P is a permutation. For non active P-boxes, the probability is obviously 1. For active P-boxes (there are $\#\Omega$ of it), we upper bound the probability by

$$\mathrm{DP}^P_{\max} = \max_{a\neq 0, b} \mathrm{DP}^P(a, b)$$

which is equal to 2^{-4} for CS-Cipher by construction. □

We thus need to lower bound $\#\Omega$ for consistent differential characteristics. This paradigm is already well known in the literature: in order to protect against heuristic differential attacks, we need to make sure that all consistent characteristic have a large number of active nonlinear boxes. Actually, the original papers of Biham and Shamir focus on looking for differential characteristics with a minimal number of S-boxes (see [7]).

Thanks to the multipermutation property of μ, it is fairly easy to investigate the minimal number of active P-boxes in CS-Cipher. Actually, μ has the property that

1. μ is a permutation,
2. for all a, both outputs of $\mu(a, y)$ are permutations of y,

3. for all b, both outputs of $\mu(x,b)$ are permutations of x.

Thus, if exactly one input of μ is non-zero, then both outputs of μ are non-zero. If the two inputs of μ are non-zero, then at least one output of μ is non-zero. In other terms, the difference patterns around one μ box can only be one out of the six following patterns.

$$00 \to 00 \quad 0* \to ** \quad *0 \to ** \quad ** \to 0* \quad ** \to *0 \quad ** \to ** .$$

(Stars mean any non-zero differences.) With the notations of the previous section, we recall that

$$\delta_{i+1} = (\mu^4 \circ L_\pi)(\delta'_i).$$

Moreover we consider consistent characteristics, which means that $(\delta'_i)_j$ is non-zero if and only if $(\delta_i)_j$ is non-zero. This enables to make rules for "non-zeroness" of the $(\delta_i)_j$.

Actually, we consider 8-bit vectors $I_i = \mathrm{Supp}(\delta_i)$. From the previous arguments we can make a list of possible $I_i \to I_{i+1}$ transitions. (In total we have $6^4 = 1296$ rules.) To each possible I_i we associate its Hamming weight $\#I_i$. We can now make the graph of all possible I_is weighted by $\#I_i$ and in which each edge corresponds to a rule. Since $\#I_i$ is also equal to the number of non-zero entries in ω_i, finding out the minimal number of active P-boxes in a consistent differential characteristic corresponds to finding a path of length 24 edges with minimal non-zero weight in this graph, which is fairly easy, for instance by using the Floyd-Warshall algorithm [14] (see [10, pp. 558–565]). Its complexity is essentially cubic in time and quadratic in memory (in term of number of vertices, which is 256 here). Experiment shows that such a path has at least a total weight of 72. More precisely, the shortest non-zero weight for paths of given edge-length is given by the table below.

l	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
w	1	3	5	9	13	18	20	24	26	30	32	36	38	42	44	48	50	54	56	60	62	66	68	72

Thus, we obtain that for any differential characteristic Ω we have

$$\mathrm{EDP}(\Omega) \leq 2^{-288}.$$

This makes CSC* provably resistant against the original differential cryptanalysis. Actually, six rounds of CSC* leads to an upper bound of 2^{-72}, which is already enough. This corresponds to two rounds of CS-Cipher instead of eight, so this suggests that **four rounds of CS-Cipher** are already secure against 2R differential attacks.

3.3 Linear Cryptanalysis

Linear cryptanalysis has a very similar treatment as was shown by Biham [6]. Actually, we have

$$\omega_{i-1} \cdot M_{i-1} = ({}^t(\mu^4)^{-1}(\omega_{i-1}) \cdot (\mu^4 \circ s_{i-1})(M_{i-1})) \oplus (\omega_{i-1} \cdot k_{i-1})$$

and

$$\omega_i \cdot M_i = {}^tL_\pi(\omega_{i+1}) \cdot (L_\pi)^{-1}(M_i)$$

thus we let

$$\delta_i = {}^t(\mu^4)^{-1}(\omega_{i-1}) \tag{15}$$

and

$$\delta'_i = {}^tL_\pi(\omega_i). \tag{16}$$

As for differential cryptanalysis, we have the following result.

Definition 7. *For any linear characteristic $\Omega = (\omega_0, \ldots, \omega_{24})$, we define the δ_is and δ'_is by Equations (15) and (16) respectively. We say that Ω is "consistent" if the property of Equation (14) holds. Let $\#\Omega$ denotes the number of indices i, j such that $(\delta_i)_j \neq 0$.*

We thus have the following result.

Lemma 8. *For any non-zero linear characteristic Ω we have*

$$\mathrm{ELP}(\Omega) \leq (\mathrm{LP}^P_{\max})^{\#\Omega}$$

where $\mathrm{LP}^P_{\max} = 2^{-4}$ *and* $\mathrm{ELP}(\Omega) = 0$ *if Ω is not consistent.*

Here the relation between the δ_is and the δ'_is is

$$\delta'_{i-1} = {}^t(\mu^4 \circ L_\pi)(\delta_i)$$

which is equivalent to

$$\delta_i = (({}^t\mu^{-1})^4 \circ L_\pi)(\delta'_{i-1})$$

instead of

$$\delta_i = (\mu^4 \circ L_\pi)(\delta'_{i-1})$$

as for differential cryptanalysis. Obviously, ${}^t\mu^{-1}$ has the same multipermutation property than μ, thus the "non-zeroness" rules for the δ_is and δ'_is are the same. We thus obtain that

$$\mathrm{ELP}(\Omega) \leq 2^{-288}.$$

This makes CSC* heuristically resistant against the original linear cryptanalysis.

3.4 Support Characteristics

Here we aim to upper bound the probabilities of the support characteristics for CSC*. One problem is that the propagation of non-zero differences through the P-boxes has no unusual cases. For this we concentrates on unusual propagations through the μ boxes.

Here the characteristic $\Omega = (I_0, \ldots, I_{24})$ defines exactly which inputs and outputs of the μ-boxes are non-zero. The probability of the characteristic is non-zero only if the number of non-zero input-output of any μ-box is in $\{0, 3, 4\}$. With these rules we can make the graph of all possible $I_i \to I_{i+1}$ transitions. The problem is to weight it.

The more interesting probabilities correspond to the case where two inputs of a μ-box are non-zero, and one output is zero. Let us denote μ_1 and μ_2 the two outputs. We thus consider the two probabilities

$$\Pr_{X,X',Y,Y'}[\mu_j(X \oplus X', Y \oplus Y') = 0 / X \neq X', Y \neq Y']$$

for $j = 1$ and $j = 2$. Due to the linearity of μ, this is equal to

$$\Pr_{X,Y}[\mu_j(X, Y) = 0 / X \neq 0, Y \neq 0]$$

which is 2^{-8} from the multipermutation properties of μ.

One problem is the vertex $I = \{1, \ldots, 8\}$, because all transitions towards it have weight 0. Intuitively, if we go through this vertex, we loose all information, and the final probability is actually meaningless, because it is smaller than the probability of the same external characteristic for a truly random cipher. For instance, if we have the support characteristic $\Omega = (I, \ldots, I_{24})$ in which $I_i = \{1, \ldots, 8\}$, we obtain

$$\Pr_{M_0, M_0', k}[E_\Omega] \leq \Pr_X[\mathrm{Supp}(X) = I_{24}]$$

so the "signal" of the support characteristic will vanish against the "noise" of natural behavior. Thus we remove this vertex from the graph.

We can now weight the edges of the graph by the total number of 2-1 transitions in the four μ-boxes: each $I_i \rightarrow I_{i+1}$ edge defines four transitions though a μ-box, so we can count the ones with no zero-input and one zero-output. Then we can look for the path with length 24 and minimal weight. The experiment shows that the minimal weight is 22. Actually, for any length $\ell \geq 2$, the minimal weight is $\ell - 2$ is obtained, for instance, by iterating the path

$$\{1, 3\} \rightarrow \{1, 2, 5, 6\} \rightarrow \{1, 3\}$$

which has weight 2 (thus probability 2^{-16}). For instance, the path

$$\{1, 3\} \rightarrow \{1, 2, 5, 6\} \rightarrow \{1, 3\} \rightarrow \ldots \rightarrow \{1, 2, 5, 6\} \rightarrow \{1, 3, 6, 8\}$$

of even length ℓ has weight $\ell - 2$.

Thus the probability of a support characteristic on 24 rounds is less than $(2^{-8})^{22} = 2^{-176}$. Actually ten rounds of CSC* leads to an upper bound of 2^{-64}. This corresponds to 3.33 rounds of CS-Cipher. So we believe that these properties make 5.33 **rounds of CS-Cipher** heuristically resistant against any multi-path differential characteristics. Eight rounds is therefore a comfortable safety margin.

One open problem is the resistance against the recent impossible differential cryptanalysis. In this paper we investigated differential characteristics with overwhelming behavior. The question now is how to address characteristics with unexpected low probabilities.

4 Conclusion

We have shown that CSC* admits no differential or linear characteristic with average probability greater than 2^{-288}, and no support characteristic with an average probability greater than 2^{-176}. We believe that these results hold for CS-Cipher as well, which makes it heuristically secure against differential, linear, truncated, and other related differential cryptanalysis. The question on the impossible differentials issue remains open though, as well as more general attacks.

Whereas ciphers similar than CS-Cipher use linear diffusion layers for mixing all pieces of a message in each round (for instance the four AES candidates Safer+, Serpent, Rijndael, Crypton), CS-Cipher uses a nonlinear diffusion primitive: the μ operation which is mixed with two non linear P-boxes. This enables to achieve a stronger design at a minimal cost (both μ and P have quite efficient implementations). It also illustrates that we can use general multipermutations and not only MDS codes: large linear layers are nice presents for the attacker.

References

1. *FIPS 46*, Data Encryption Standard. U.S. Department of Commerce — National Bureau of Standards, National Technical Information Service, Springfield, Virginia. *Federal Information Processing Standard Publication 46*, 1977.
2. *CD-ROM "AES CD-1: Documentation"*, National Institute of Standards and Technology (NIST), August 1998. Documentation for the First Advanced Encryption Standard Candidate Conference.
3. C. M. Adams. *A Formal and Practical Design Procedure for Substitution-Permutation Network Cryptosystems.*, Ph.D. Thesis of Queen's University, Kingston, Ontario, Canada, 1990.
4. C. M. Adams, S. E. Tavares. Designing s-boxes Resistant to Differential Cryptanalysis. In *Proceedings of 3rd Symposium on the State and Progress of Research in Cryptography*, pp. 386–397, Rome, Italy, 1994.
5. O. Baudron, H. Gilbert, L. Granboulan, H. Handschuh, R. Harley, A. Joux, P. Nguyen, F. Noilhan, D. Pointcheval, T. Pornin, G. Poupard, J. Stern, S. Vaudenay. DFC Update. In Proceedings from the Second Advanced Encryption Standard Candidate Conference, National Institute of Standards and Technology (NIST), March 1999.
6. E. Biham. On Matsui's Linear Cryptanalysis. In *Advances in Cryptology EUROCRYPT'94*, Perugia, Italy, Lectures Notes in Computer Science 950, pp. 341–355, Springer-Verlag, 1995.
7. 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.
8. E. Biham, A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.
9. 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.

10. T. H. Cormen, C. E. Leiserson, R. L. Rivest. *Introduction to Algorithms*, Mc Graw Hill, 1990.
11. J. Daemen. *Cipher and Hash Function Design — Strategies based on Linear and Differential Cryptanalysis*, Doctoral Dissertation, Katholieke Universiteit Leuven, 1995.
12. J. Daemen, L. R. Knudsen, V. Rijmen. The Block Cipher Square. In *Fast Software Encryption*, Haifa, Israel, Lectures Notes in Computer Science 1267, pp. 149–165, Springer-Verlag, 1997.
13. H. Feistel. Cryptography and Computer Privacy. *Scientific American*, vol. 228, pp. 15–23, 1973.
14. R. W. Floyd. Algorithm 97 (SHORTEST PATH). In *Communications of the ACM*, vol. 5, p. 345, 1962.
15. H. Gilbert, M. Girault, P. Hoogvorst, F. Noilhan, T. Pornin, G. Poupard, J. Stern, S. Vaudenay. Decorrelated Fast Cipher: an AES Candidate. Submitted to the Advanced Encryption Standard process. In *CD-ROM "AES CD-1: Documentation"*, National Institute of Standards and Technology (NIST), August 1998.
16. H. M. Heys. *The Design of Substitution-Permutation Network Ciphers Resistant to Cryptanalysis*, Ph.D. Thesis of Queen's University, Kingston, Ontario, Canada, 1994.
17. H. M. Heys, S. E. Tavares. Substitution-Permutation Networks Resistant to Differential and Linear Cryptanalysis. *Journal of Cryptology*, vol. 9, pp. 1–19, 1996.
18. L. R. Knudsen. Truncated and Higher Order Differentials. In *Fast Software Encryption*, Leuven, Belgium, Lectures Notes in Computer Science 1008, pp. 196–211, Springer-Verlag, 1995.
19. L. R. Knudsen, T. A. Berson. Truncated Differentials of SAFER. In *Fast Software Encryption*, Cambridge, United Kingdom, Lectures Notes in Computer Science 1039, pp. 15–26, Springer-Verlag, 1996.
20. X. Lai. *On the Design and Security of Block Ciphers*, ETH Series in Information Processing, vol. 1, Hartung-Gorre Verlag Konstanz, 1992.
21. 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.
22. J. L. Massey. SAFER K-64: a Byte-Oriented Block-Ciphering Algorithm. In *Fast Software Encryption*, Cambridge, United Kingdom, Lectures Notes in Computer Science 809, pp. 1–17, Springer-Verlag, 1994.
23. 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.
24. 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.
25. 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.
26. M. Matsui. New Block Encryption Algorithm MISTY. In *Fast Software Encryption*, Haifa, Israel, Lectures Notes in Computer Science 1267, pp. 54–68, Springer-Verlag, 1997.
27. 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.

28. K. Nyberg. Differentially Uniform Mapping for Cryptography. In *Advances in Cryptology EUROCRYPT'93*, Lofthus, Norway, Lectures Notes in Computer Science 765, pp. 55–64, Springer-Verlag, 1994.
29. K. Nyberg. Linear Approximation of Block Ciphers. In *Advances in Cryptology EUROCRYPT'94*, Perugia, Italy, Lectures Notes in Computer Science 950, pp. 439–444, Springer-Verlag, 1995.
30. K. Nyberg, L. R. Knudsen. Provable Security against a Differential Cryptanalysis. In *Advances in Cryptology CRYPTO'92*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 740, pp. 566–574, Springer-Verlag, 1993.
31. K. Nyberg, L. R. Knudsen. Provable Security against a Differential Cryptanalysis. *Journal of Cryptology*, vol. 8, pp. 27–37, 1995.
32. J. Stern, S. Vaudenay. CS-Cipher. In *Fast Software Encryption*, Paris, France, Lectures Notes in Computer Science 1372, pp. 189–205, Springer-Verlag, 1998.
33. S. Vaudenay. On the Need for Multipermutations: Cryptanalysis of MD4 and SAFER. In *Fast Software Encryption*, Leuven, Belgium, Lectures Notes in Computer Science 1008, pp. 286–297, Springer-Verlag, 1995.
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. A cheap Paradigm for Block Cipher Security Strengthening. Technical Report LIENS-97-3, 1997.
36. 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.
37. S. Vaudenay. Feistel Ciphers with L_2-Decorrelation. To appear in SAC'98, LNCS.
38. S. Vaudenay. The Decorrelation Technique Home-Page. URL:`http://www.dmi.ens.fr/~vaudenay/decorrelation.html`
39. S. Vaudenay. Resistance Against General Iterated Attacks. (To appear in Eurocrypt' 99.)
40. A. M. Youssef, S. Mister, S. E. Tavares. On the Design of Linear Transformations for Substitution Permutation Encryption Networks. Presented at SAC'97.