On the Distribution of Characteristics in Bijective Mappings

Luke O'Connor [1]
University of Waterloo, Canada

Abstract

Differential cryptanalysis is a method of attacking iterated mappings which has been applied with varying success to a number of product ciphers and hash functions [1, 3]. The attack is based on predicting a series of differences $\Delta Y_1, \Delta Y_2, \ldots, \Delta Y_r$ known as a *characteristic* Ω. Partial information about the key can be derived when the differences are correctly predicted. The probability of a given characteristic Ω correctly predicting differences is derived from the XOR tables associated with the iterated mapping.

Even though differential cryptanalysis has been applied successfully to a number of specific iterated mappings such as DES, FEAL and LOKI, the effectiveness of the attack against an arbitrary iterated mapping has not been considered. In this paper we derive the exact distribution of characteristics in XOR tables, and determine an upper bound on the probability of the most likely characteristic Ω in a product cipher constructed from randomly selected S-boxes that are bijective mappings. From this upper bound we are then able to construct product ciphers for which all characteristics Ω occur with low probability.

Keywords: Differential cryptanalysis, iterated mapping, product cipher.

1 Introduction and Results

Differential cryptanalysis is a statistical attack popularized by Biham and Shamir in a series of well-known papers [1, 2, 3]. The attack has been applied to a wide range of iterated mappings including LUCIFER, DES, FEAL, REDOC, Kahfre [4, 5, 12, 13, 17, 19]. As explained below, the attack is based on a quantity Ω called a *characteristic*, which has some probability p^Ω of giving information about the secret key used in the mapping. The attack is universal in that characteristics Ω will always exist for any iterated mapping;

[1] The current employer of the author is the Distributed System Technology Center (DSTC), Brisbane, Australia. Correspondence should be sent to ISRC, QUT Gardens Point, 2 George Street, GPO Box 2434, Brisbane Q 4001, Australia: email oconnor@sleet.fit.qut.edu.au.

T. Helleseth (Ed.): Advances in Cryptology - EUROCRYPT '93, LNCS 765, pp. 360-370, 1994.
© Springer-Verlag Berlin Heidelberg 1994

however p^{Ω} may be very small, and possibly less likely to furnish information concerning the key than the success of guessing the secret key at random. For this reason, differential cryptanalysis has had varying success against the iterated mappings it has been applied to, and little is known about how useful the attack is expected to be against an arbitrary iterated mapping.

In Figure 1 we present the basic substitution-transposition network (ST-network) [6]: each round consists of several small substitutions S (S-boxes) followed by a transposition T (anagram) of the current ciphertext. This model is generally acknowledged [6, 9, 18] as being a practical realization of product ciphers originally proposed by Shannon [16]. Most product ciphers such as LUCIFER, DES and IDEA are variations or extensions of the basic ST-network. The main result of this paper is to determine how well differential cryptanalysis is expected to perform against randomly generated instances of ST-networks.

We will give a brief description of differential cryptanalysis with reference to product ciphers, though any iterated mapping would suffice. For a product cipher E that consists of R rounds, let $E_r(X, K)$ be the encryption of the plaintext X under the key K for r rounds, $1 \leq r \leq R$. Note that $E_R(X, K) = E(X, K) = C$ is the ciphertext for X. Let $\Delta C(r) = E_r(X, K) + E_r(X', K)$ be the difference between the ciphertexts of two plaintexts X, X' after r rounds where $1 \leq r \leq R$. For our purposes the difference operator $+$ will refer to addition in the vector space Z_2^m. An r-round *characteristic* is defined as an $(r+1)$-tuple $\Omega_r(\Delta X, \Delta Y_1, \Delta Y_2, \ldots, \Delta Y_r)$ where ΔX is a plaintext difference, and the ΔY_i are ciphertext differences. A plaintext pair X, X' of difference $\Delta X = X + X'$ is called

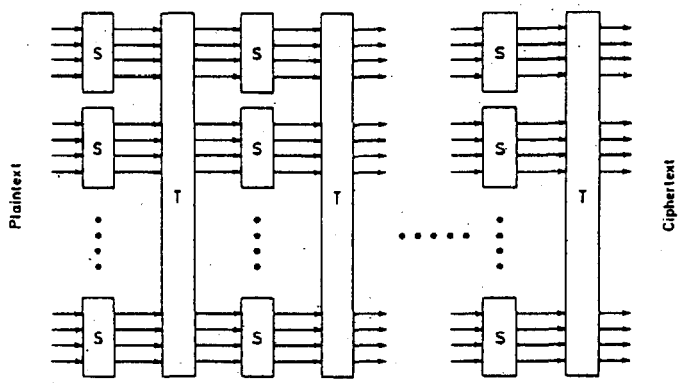

Figure 1: The ST-network product cipher

a *right pair* with respect to a key K and a characteristic $\Omega_r(\Delta X, \Delta Y_1, \Delta Y_2, \ldots, \Delta Y_r)$ if when X and X' are encrypted, $\Delta C(i) = \Delta Y_i$ for $1 \leq i \leq r$. That is, X, X' is a right pair if the characteristic correctly predicts the ciphertext differences at each round. The characteristic Ω_r has probability p^{Ω_r} if a fraction p^{Ω_r} of the plaintext pairs of difference

ΔX are right pairs. On the other hand, if X, X' such that $\Delta X = X + X'$ is not a right pair, then it is said to be a *wrong pair* (with respect to the characteristic and the key). A table which records the number of pairs of difference ΔX that give the output difference ΔY for a mapping π is called the *XOR table distribution* of π. A characteristic $\Delta X, \Delta Y$ is said to be *impossible* for π if its corresponding XOR table entry is zero. Also a characteristic will be called *nonzero* if $w(\Delta X), w(\Delta Y) > 0$, where $w(\cdot)$ is the Hamming weight function. Using a characteristic Ω of appropriate length it is then possible to devise a statistical experiment which when repeated a sufficient number of times will yield the subkey of the last round (see [1] for details).

Product ciphers such as LUCIFER, DES, FEAL and IDEA are iterated mappings that use a fixed mapping G at each round. For example, in DES the function G at round i, $1 \leq i \leq 16$ is defined as

$$L_i \circ R_i \;=\; G(L_{i-1} \circ R_{i-1}) \;=\; R_{i-1} \circ [\, L_{i-1} + P(S(E(R_{i-1}) + K_i)) \,]$$

where $\circ$ denotes string concatenation, E is a 32-to-48-bit expansion, S is a substitution by 8×6-to-4-bit S-boxes, and P is a 32-element transposition. When the components of G are fixed, which for DES are the E, S and P mappings, we observe that an r-round characteristic is simply the concatenation of r 1-round, or single round, characteristics defined on the mapping G. For an r-round characteristic $\Omega_r(\Delta X, \Delta Y_1, \Delta Y_2, \ldots, \Delta Y_r)$ we have

$$p^{\Omega_r} \;=\; \Pr(\Delta C(i) = \Delta Y_i,\;\; 1 \leq i \leq r \mid X + X' = \Delta X) \;\leq\; \prod_{i=0}^{r-1} p^{\omega_i} \qquad (1)$$

where $\Delta Y_0 = \Delta X$ and ω_i is the single round characteristic $\Delta Y_i, \Delta Y_{i+1}$ for $0 \leq i \leq r-1$ defined on G. It then follows that the probability of the r-round characteristic Ω_r can be bound as $p^{\Omega_r} \leq (p^{\Omega})^r$ where p^{Ω} is the *probability of the most likely (nonzero) single round characteristic.*

At present there are no general bounds known for p^{Ω}; indeed it is difficult to give a definition of a 'general' iterated mapping which can be used for deriving bounds on p^{Ω}. What can be said with certainty is that a product cipher E which claims to be useful must be bijective (plaintexts are taken to distinct ciphertexts). This suggests that the XOR properties of bijective mappings should be examined. If this examination is successful, then we may apply these results to the ST-networks of Figure 1 where the S-boxes themselves are bijective so as to ensure that the mapping itself is bijective.

Let $\pi : Z_2^m \rightarrow Z_2^m$ be a bijective mapping, referred to as an m-bit permutation. The set of all m-bit permutations is known as the symmetric group on 2^m objects and

is denoted as S_{2^m}. Let $\Lambda_\pi(\Delta X, \Delta Y)$ be the value of the XOR table entry of the pair $\Delta X, \Delta Y \in Z_2^m$ for the permutation $\pi \in S_{2^m}$. Assuming the uniform distribution on the set S_{2^m} we prove (Theorem 2.1) that

$$\Pr(\Lambda_\pi(\Delta X, \Delta Y) = 0) \;=\; \frac{1}{2^m!} \cdot \sum_{k=0}^{2^{m-1}} (-1)^k \cdot \binom{2^{m-1}}{k}^2 \cdot 2^k \cdot k! \cdot (2^m - 2k)!.$$

We are then able to show (Corollary 3.1) that for large m, the expected probability of the most likely nonzero characteristic for an m-bit permutation is at most $\frac{m}{2^{m-1}}$. Equivalently, the expected maximum entry in the XOR table for nonzero characteristics is at most $2m$ for large m. The result of Corollary 3.1 can used to estimate the probability of the most likely 1-round characteristic p^Ω in an iterated mapping based on m-bit permutations. Consider a 16-round 64-bit product cipher E for which the round mapping consists of $8 \times$ 8-bit permutations followed by a 64-bit transposition. Then to predict the input difference to the 16th round requires a 15-round characteristic Ω_{15} where the input difference to each of the first 15 rounds is nonzero. Let us assume that the permutations are selected uniformly from S_{2^8} and that at each round there is only one S-box which has a nonzero input difference. It then follows from Corollary 3.1 that

$$p^{\Omega_{15}} \;\le\; \left(p^\Omega\right)^{15} \;\le\; \left(\frac{8}{2^7}\right)^{15} \;=\; 0.86736 \times 10^{-18} \;\approx\; 2^{-59}.$$

Further, if Ω_{15} has nonzero input differences to two S-boxes at 7 out of the 15 rounds then

$$p^{\Omega_{15}} \;\le\; \left(\frac{8}{2^7}\right)^{2 \cdot 7} \cdot \left(\frac{8}{2^7}\right)^{8} \;=\; 0.32311 \times 10^{-26} \;\approx\; 2^{-86}.$$

Corollary 3.1 indicates that the individual entries of an XOR table are expected to be distributed in the interval $[0, 2, \ldots, 2m]$. At this point we are not able to determine the exact distribution of entries within this interval, but we are able to prove that most XOR table entries are in fact zero. We prove (Theorem 3.2) that the expected fraction of the XOR table for nonzero characteristics that is zero approaches $e^{-\frac{1}{2}} = 0.60653$. In another way, approximately 60% of the entries for nonzero characteristics will be zero for a permutation selected uniformly.

The full proofs of the theorems to follow are ommitted since the final version of this paper has been accepted for publication in the *Journal of Cryptology*.

1.1 Notation

We will now formalize some of the notation given in the introduction. Let $[\cdot]$ be a boolean predicate that evaluates to 0 or 1 such as [n is prime]. For a given $\pi \in S_{2^m}$, define $\Lambda_\pi(\Delta X, \Delta Y)$ as

$$\Lambda_\pi(\Delta X, \Delta Y) = \sum_{\substack{X, X' \in Z_2^m \\ \Delta X = X + X'}} [\pi(X) + \pi(X') = \Delta Y]. \quad (2)$$

Thus $2^{-m} \cdot \Lambda_\pi(\Delta X, \Delta Y)$ is a random variable giving the probability that the difference in the output of the mapping π is ΔY when the difference of the input pair X, X' is ΔX. For all $\pi \in S_{2^m}$, observe that when $\Delta X = 0$ or $\Delta Y = 0$ it follows that $\Lambda_\pi(\Delta X, \Delta Y) = 0$, unless $\Delta X = \Delta Y = 0$ whereupon $\Lambda_\pi(\Delta X, \Delta Y) = 2^m$. The distribution of $\Lambda_\pi(\Delta X, \Delta Y)$ taken over all possible $\Delta X, \Delta Y \in Z_2^m$ is known as the *pairs XOR distribution table for* π, or simply the XOR table for π.

Example 1.1 For an m-bit permutation π, let XOR_π be the $2^m \times 2^m$ matrix where $\mathrm{XOR}_\pi(i,j) = \Lambda_\pi(i,j)$, $0 \le i, j \le 2^m - 1$, where i, j are treated as 3-bit binary vectors. Observe that $\mathrm{XOR}_\pi(0,0) = 8$, and all other entries in the first row or column of $\mathrm{XOR}(\pi)$ are zero. For $\pi = (7,2,4,1,5,6,3,0)$ the corresponding XOR table is given as:

$$\mathrm{XOR}_\pi = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 4 & 0 \\ 0 & 2 & 2 & 0 & 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 & 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 & 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 & 2 & 0 & 0 & 2 \end{bmatrix}. \quad (3)$$

Notice that if each entry in the XOR table is divided by 2^m then the resulting matrix will be doubly stochastic. □

The XOR table for an m-bit permutation π has the following general form:

$$\mathrm{XOR}_\pi = \begin{bmatrix} 2^m & 0 & 0 & \cdots & 0 \\ 0 & a_{1,1} & a_{1,2} & \cdots & a_{1,2^m-1} \\ 0 & a_{2,1} & a_{2,2} & \cdots & a_{2,2^m-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & a_{2^m-1,1} & a_{2^m-1,2} & \cdots & a_{2^m-1,2^m-1} \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} 2^m & 0 \\ 0 & A_\pi \end{bmatrix}. \quad (4)$$

We are interested in the properties of the $(2^m - 1) \times (2^m - 1)$ submatrix $A_\pi = [a_{i,j}]$, $1 \le i, j \le 2^m - 1$, which corresponds to that portion of the XOR table entries attributed to nonzero characteristics. In this paper we will show that for large m, approximately 60% of the entries in A_π are zero and largest entry in A_π is expected to be bounded by $2m$.

2 The Pairing Theorem

Observe that a characteristic $\Delta X, \Delta Y$ corresponds to a pairing of the inputs and outputs of a permutation π (namely the pairs X, X' and Y, Y' where $\Delta X = X + X'$ and $\Delta Y = Y + Y'$). For $\phi : A \to B$, let Π_A and Π_B be pairings on the sets A and B, respectively. Theorem 2.1 determines the number of functions ϕ which take no pair of Π_A to a pair in Π_B, and will be referred to as the Pairing Theorem.

Theorem 2.1 (Pairing Theorem) Let $A = \{a_1, a_2, \ldots, a_{2d}\}$ and $B = \{b_1, b_2, \ldots, b_{2d}\}$ be sets of distinct elements. Let $\Pi_A \subseteq A \times A$ and $\Pi_B \subseteq B \times B$ be unordered pairs, such that $a_i(b_i)$ occurs in one pair of $\Pi_A(\Pi_B)$ for $1 \leq i \leq 2d$. Then the number $\Phi(d)$ of bijective functions $\phi : A \to B$ such that for $\forall(a_i, a_j) \in \Pi_A$, $(\phi(a_i), \phi(a_j)) \notin \Pi_B$ is

$$\Phi(d) = \sum_{k=0}^{d} (-1)^k \cdot \binom{d}{k}^2 \cdot 2^k \cdot k! \cdot (2d-2k)!. \tag{5}$$

Proof. Order the elements of Π_B as (b'_i, b'_{d+i}), $1 \leq i \leq d$. For $1 \leq i \leq d$ define $P(i)$ as

$$P(i) = \{\ \phi \mid (\phi(a), \phi(a')) = (b'_i, b'_{d+i}),\ (a, a') \in \Pi_A\ \}$$

which is the number of functions ϕ that map some pair of Π_A to the pair $(b'_i, b'_{d+i}) \in \Pi_B$. It follows that

$$\Phi(d) = (2d)! - \left| \bigcup_{1 \leq j \leq d} P(j) \right| = (2d)! + \sum_{\substack{S \subseteq \{1,2,\ldots,d\} \\ S \neq \{\emptyset\}}} (-1)^{|S|} \cdot \left| \bigcap_{j \in S} P(j) \right| \tag{6}$$

using the inclusion-exclusion principle [8]. For $1 \leq k \leq d$ define the integers

$$P(i'_1, i'_2, \ldots, i'_k) = \left| \bigcap_{1 \leq j \leq k} P(i'_j) \right| \tag{7}$$

and by symmetry $P(1, 2, \ldots, k) = P(i'_1, i'_2, \ldots, i'_k) \stackrel{\text{def}}{=} P(d, k)$. From eq. (6) it then follows that

$$\Phi(d) = (2d)! + \sum_{k=1}^{d} (-1)^k \cdot \binom{d}{k} \cdot P(d, k). \tag{8}$$

It remains to determine $P(d, k)$ for $1 \leq k \leq d$. To this end, order the pairs within Π_A as (a'_i, a'_{d+i}) for $1 \leq k \leq d$. Then $P(d, k)$ is the number of functions ϕ for which there are k pairs (a''_i, a''_{d+i}) such that $\{\phi(a''_i), \phi(a''_{d+i})\} = \{(b'_i, b'_{d+i})\}$, $1 \leq i \leq k$. There are $\binom{d}{k}$ ways to choose the k pairs (a''_i, a''_{d+i}) from Π_A, $k!$ ways of assigning the (a''_i, a''_{d+i}) to the (b'_i, b'_{d+i}), and 2^k ways of assigning (a''_i, a''_{d+i}) to a particular pair in Π_B. It then follows that

$$P(d, k) = \binom{d}{k} \cdot 2^k \cdot k! \cdot (2d-2k)! \tag{9}$$

where $(2d-2k)!$ is the number of ways to assign the elements in $A-\{a''_i, a''_{d+i} \mid 1 \le i \le k\}$. We then have that

$$\Phi(d) \;=\; (2d)! + \sum_{k=1}^{d} (-1)^k \cdot \binom{d}{k} \cdot P(d,k) \;=\; \sum_{k=0}^{d} (-1)^k \cdot \binom{d}{k}^2 \cdot 2^k \cdot k! \cdot (2d-2k)!$$

which completes the proof of the theorem. □

It is a simple matter to observe that for a fixed mapping π the expected value of each entry in the XOR table is 1 since there are 2^m entries in each row which sum to 2^m. Using the Pairing Theorem we are now able to derive the exact distribution of the random variable $\Lambda_\pi(\Delta X, \Delta Y)$.

Corollary 2.1 For any fixed nonzero $\Delta X, \Delta Y \in Z_2^m$, assuming π is chosen uniformly from the set S_{2^m}, and $0 \le k \le 2^{m-1}$

$$\Pr(\Lambda_\pi(\Delta X, \Delta Y) = 2k) \;=\; \binom{2^{m-1}}{k}^2 \cdot \frac{k! \cdot 2^k \cdot \Phi(2^{m-1}-k)}{2^m!}. \tag{10}$$

□

3 Two properties of XOR tables

Recall that p^Ω was defined in the introduction as the probability of the most likely single round characteristic for an iterated mapping. In this section we will derive bounds on the expected value of p^Ω assuming that the round mapping G is based on m-bit permutations selected uniformly from S_{2^m}. Let G consist of s S-boxes implementing m-bit permutations $\pi_1, \pi_2, \ldots, \pi_s$ such that $G : Z_2^{m \cdot s} \rightarrow Z_2^{m \cdot s}$ where π_1 operates on the first block of s bits, π_2 operates on the second block of s bits, and so on, as in Figure 1. For example, G may operate on 48 bits of ciphertext which is used as the input to 8×6-bit permutations $\pi_1, \pi_2, \ldots, \pi_8$. Then define Λ_m^* as

$$2^m \cdot p^\Omega \;\le\; \Lambda_m^* \;\stackrel{\text{def}}{=}\; \max_{\substack{\pi \in \{\pi_1, \pi_2, \ldots, \pi_s\} \\ \Delta X, \Delta Y \in Z_2^m \\ \omega(\Delta X), \omega(\Delta Y) > 0}} \Lambda_\pi(\Delta X, \Delta Y)$$

from which it follows that $\frac{\Lambda_m^*}{2^m}$ is the probability of the most likely characteristic across all s permutations in G. Then for any nonzero r-round characteristic Ω_r it follows that

$$p^{\Omega_r} \;\le\; \left(\frac{\Lambda_m^*}{2^{m-1}}\right)^r. \tag{11}$$

At present there are no known general bounds on Λ_m^*. For a randomly selected set of

m-bit permutations $\{\pi_1, \pi_2, \ldots, \pi_s\}$ we may use the Pairing Theorem to determine an expected upper bound on Λ_m^*.

Theorem 3.1 Assuming that the S-boxes π_i are selected uniformaly from the set S_{2^m}

$$\lim_{m\to\infty} \frac{\mathbf{E}[\Lambda_m^*]}{2m} \leq 1. \tag{12}$$

□

Sketch of proof. For $1 \leq k \leq 2^{m-1}$, let $\Lambda_{m,2k}$ be the expected number of nonzero characteristics $\Delta X, \Delta Y$ for which $\Lambda_\pi(\Delta X, \Delta Y) = 2k$. Further let $\Pr(\Lambda_\pi = 2k)$ be the probability that an m-bit permutation has a nonzero characteristic $\Delta X, \Delta Y$ for which $\Lambda_\pi(\Delta X, \Delta Y) = 2k$. The proof rests on the following inequality:

$$\Pr(\Lambda_m^* = 2k) \;<\; \Pr(\Lambda_\pi = 2k) \;\leq\; \Lambda_{m,2k}.$$

Using the Pairing Theorem it can be shown that

$$\Lambda_{m,2k} \;=\; \frac{(2^m-1)^2}{2^m!} \cdot \binom{2^{m-1}}{k}^2 \cdot 2^k \cdot k! \cdot \Phi(2^{m-1}-k). \tag{13}$$

The theorem follows from proving that $\sum_{k>m} 2k \cdot \Lambda_{m,2k} = o(1)$. □

Corollary 3.1 For large m and assuming the uniform distribution on the set S_{2^m}, the expected probability of the most likely nonzero characteristic is bounded by $\frac{m}{2^{m-1}}$.

Proof. The expected probability of the most likely nonzero characteristic is $\frac{\Lambda_m^*}{2^m}$. □

Let $\overline{\Lambda_m} = \sum_{k=m+1}^{2^{m-1}} 2k \cdot \Lambda_{m,2k}$ be an upper bound on the last $2^{m-1} - m$ terms in the sum for $\mathbf{E}[\Lambda_m^*]$. Also let $\overline{\Lambda_m^*}$ be an empirical estimate of $\mathbf{E}[\Lambda_m^*]$ based on a sample of m_p random permutations. Further, let min (max) be the smallest (largest) maximum XOR entry found across the m_p permutations. Table 1 lists these quantities for $m = 4, 5, \ldots, 10$. We see that $2(m+1) \cdot \Lambda_{m,2(m+1)}$ is a good approximation of $\overline{\Lambda_m^*}$, and by $m = 6$ the tail of the summation for $\mathbf{E}[\Lambda_m^*]$ beginning at $k = 2(m+1)$ is negligibly small.

The presence of impossible characteristics assists in discarding certain plaintext pairs which cannot give any probabilistic information concerning the key. It has been observed that 20% –30% of the characteristics in the S-boxes of DES are impossible. Let $\Lambda_{m,0}$ be the expected number of nonzero characteristics $\Delta X, \Delta Y$ which have zero entries in the XOR table of a uniformly selected m-bit permutation. We are able to compute $\Lambda_{m,0}$ as a direct application of the Pairing Theorem.

Theorem 3.2 For any fixed nonzero $\Delta X, \Delta Y \in Z_2^m$ and assuming π is chosen uniformly from the set S_{2^m}

$$\lim_{m\to\infty} \Lambda_{m,0} \;=\; \frac{(2^m-1)^2}{e^{\frac{1}{2}}}. \tag{14}$$

m	$2(m+1)\cdot \Lambda_{m,2(m+1)}$	$\overline{\Lambda_m}$	$\overline{\Lambda_m^-}$	min	max	m_p
4	.76863	.87258	3.114	2	6	10,000
5	.25973	.28436	3.839	3	6	10,000
6	$.80244 \times 10^{-1}$	$.86489 \times 10^{-1}$	4.495	3	7	10,000
7	$.22027 \times 10^{-1}$	$.23498 \times 10^{-1}$	5.126	4	8	10,000
8	$.53856 \times 10^{-2}$	$.57019 \times 10^{-2}$	5.606	5	8	1000
9	$.11818 \times 10^{-2}$	$.12438 \times 10^{-2}$	6.190	6	8	1000
10	$.23470 \times 10^{-3}$	$.24584 \times 10^{-3}$	6.700	6	9	1000

Table 1: The distribution of characteristics.

Sketch of proof. Recall from the Pairing Theorem that $\Phi(2^{m-1})$ will give the number of m-bit permutations that π for which a given characteristic $\Delta X, \Delta Y$ is impossible. It can be shown that the alternating sum in eq. (5) is dominated by its first term ($k = 0$), and that $\Phi(d) \sim (2d!)/e^{\frac{1}{2}}$. □

It now follows that approximately 60% of the entries of the A_π submatrix defined in eq. (4) are zero since $e^{-\frac{1}{2}} = 0.6065$.

4 Conclusion and Remarks

Our results then show that a relatively simple design can produce product ciphers for which all characteristics Ω are expected to (correctly) predict differences with low probability. We further note that random m-bit permutations can be generated efficiently [15], and that the fraction of permutations that are with linear [7] or degenerate [14] in any output bit is tending to zero rapidly as a function of m. On the other hand, Biham and Shamir [3] found that replacing the S-boxes of DES by random 4-bit permutations yielded systems that were far weaker than the original DES. The weakness of these S-boxes appears to be due to the dimension of the permutation rather than the use of permutations *per se*.

References

[1] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, 4(1):3–72, 1991.

[2] E. Biham and A. Shamir. Differential cryptanalysis of the full 16-round DES. Technical Report 708, Technion, Israel Institute of Technology, Haifa, Israel, 1991.

[3] E. Biham and A. Shamir. Differential cryptanalysis of Snefru, Khafre, REDOC-II, LOKI and LUCIFER. *Advances in Cryptology, CRYPTO 91, Lecture Notes in Computer Science, vol. 576, J. Feigenbaum ed., Springer-Verlag*, pages 156–171, 1992.

[4] L. P. Brown, J. Pieprzyk, and J. Seberry. LOKI - a cryptographic primitive for authentication and secrecy applications. *Advances in Cryptology, AUSCRYPT 90, Lecture Notes in Computer Science, vol. 453, J. Seberry and J. Pieprzyk eds., Springer-Verlag*, pages 229–236, 1990.

[5] T. Cusick and M. Wood. The REDOC-II cryptosystem. *Advances in Cryptology, CRYPTO 90, Lecture Notes in Computer Science, vol. 537, A. J. Menezes and S. A. Vanstone ed., Springer-Verlag*, pages 545–563, 1991.

[6] H. Feistel. Cryptography and computer privacy. *Scientific American*, 228(5):15–23, 1973.

[7] J. Gordon and H. Retkin. Are big S-boxes best? In T. Beth, editor, *Cryptography, proceedings, Burg Feuerstein*, pages 257–262, 1982.

[8] M. Hall. *Combinatorial Theory.* Blaisdell Publishing Company, 1967.

[9] J. B. Kam and G. I. Davida. A structured design of substitution-permutation encryption networks. *IEEE Transactions on Computers*, 28(10):747–753, 1979.

[10] X. Lai. *On the design and security of block ciphers.* ETH Series in Information Processing, editor J. Massey, Hartung-Gorre Verlag Konstanz, 1992.

[11] X. Lai, J. Massey, and S. Murphy. Markov ciphers and differential analysis. In *Advances in Cryptology, EUROCRYPT 91, Lecture Notes in Computer Science, vol. 547, D. W. Davies ed., Springer-Verlag*, pages 17–38, 1991.

[12] X. Lai and J. L. Massey. A proposal for a new block encryption standard. In *Advances in Cryptology, EUROCRYPT 90, Lecture Notes in Computer Science, vol. 473, I. B. Damgård ed., Springer-Verlag*, pages 389–404, 1991.

[13] R. Merkle. Fast software encryption functions. *Advances in Cryptology, CRYPTO 90, Lecture Notes in Computer Science, vol. 537, A. J. Menezes and S. A. Vanstone ed., Springer-Verlag*, pages 476–501, 1991.

[14] L. J. O'Connor. Enumerating nondegenerate permutations. *Advances in Cryptology, EUROCRYPT 91, Lecture Notes in Computer Science, vol. 547, D. W. Davies ed., Springer-Verlag*, pages 368–377, 1991.

[15] E. M. Reingold, J. Nievergeld, and N. Deo. *Combinatorial Algorithms: Theory and Practice.* Prentice-Hall, 1976.

[16] C. E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28:656–175, 1949.

[17] A. Shimizu and S. Miyaguchi. Fast data encipherment algorithm FEAL. *Advances in Cryptology, EUROCRYPT 87, Lecture Notes in Computer Science, vol. 304, D. Chaum and W. L. Price eds., Springer-Verlag*, pages 267–278, 1988.

[18] N. J. A. Sloane. Error correcting codes and cryptography, part 1. *Cryptologia*, 6(2):128–153, 1982.

[19] A. Sorkin. LUCIFER: a cryptographic algorithm. *Cryptologia*, 8(1):22–35, 1984.