On the Distribution of Characteristics in Composite Permutations

Luke O'Connor

Distributed Systems Technology Center
Brisbane, Australia
email: oconnor@fitmail.fit.qut.edu.au

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, 2]. Let $\rho : Z_2^c \times Z_2^m \rightarrow Z_2^m$ be a mapping that consists of c 'control' bits and m 'data' bits. The mapping ρ mapping contains 2^c m-bit permutations $\pi_i : Z_2^m \rightarrow Z_2^m$, $0 \leq i \leq 2^c - 1$, one of which is selected (multiplexed) by the control bits, and a substitution is then performed on the data bits using the selected permutation. Such mappings will be called *composite permutations*. The S-boxes of DES are composite permutations of the form $Si : Z_2^2 \times Z_2^4 \rightarrow Z_2^4$ with 2 control bits and 4 data bits.
In differential cryptanalysis the attacker is interested in the largest entry in a given XOR table, and the fraction of the XOR table that is zero. In this paper we determine the distribution of characteristics in the XOR tables of composite permutations, which leads to approximations for the largest entry in the XOR table and the density of zero entries.

Keywords: Differential cryptanalysis, iterated mapping, product cipher.

D.R. Stinson (Ed.): Advances in Cryptology - CRYPTO '93, LNCS 773, pp. 403-412, 1994.

1 Introduction and Results

Differential cryptanalysis is a statistical attack popularized by Biham and Shamir [1, 2] that has been applied to a wide range of iterated mappings including LUCIFER, DES, FEAL, REDOC, Kahfre [3, 4, 8, 9, 12, 13]. 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, though p^{Ω} may be very small, and possibly less likely than the probability of guessing the secret key at random.

An r-round characteristic Ω is an $(r+1)$-tuple of differences $\Delta X, \Delta Y_1, \ldots, \Delta Y_r$; the probability p^{Ω} is defined as the fraction of plaintext pairs X, X' for which $X + X' = \Delta X$ and ΔY_i is the difference[1] of the encryption of X and X' after i rounds for all $1 \leq i \leq r$. If Ω incorrectly predicts an intermediate difference ΔY_i for a plaintext pair X, X' satisfying $X + X' = \Delta X$, then X, X' is said to be a *wrong pair* with respect to the characteristic. The probability of the differences ΔY_i being correctly predicted will typically depend on the distribution of differences in the auxiliary tables used by the iterated mapping. The *XOR table* of a mapping $\rho : Z_2^n \rightarrow Z_2^m$ shows the number of input pairs of difference $\Delta X \in Z_2^n$ that map to outputs of difference $\Delta Y \in Z_2^m$. In the case of DES, these auxiliary tables are known as S-boxes, and a study of the corresponding XOR tables by Biham and Shamir [1] yielded the following information: (i) the most likely input/output difference pair for any one S-box occurs with probability $\frac{1}{4}$; (ii) approximately 70–80% of the entries in the XOR tables are zero ; (iii) it is conjectured that if Ω is an r-round characteristic then extending Ω by an additional 2 rounds will decrease p^{Ω} by a factor of at least β where $\beta \approx 1/234$.

Each S-box in DES is a union of 4 permutations of the integers $\{0, 1, \ldots, 15\}$ where 2 of the 6 input bits select the permutation that will be used as the current substitution. We will call such mappings *composite permutations*. The design criteria of employing composite permutations in DES has not been adequately explained, but the experimental work of Dawson and Tavares [5] indicates that composite permutations have better XOR table distributions than single permutations $\pi : Z_2^m \rightarrow Z_2^m$ (as well as being close to optimal with respect to several other design criteria for S-boxes). In this paper we complement this experimental work by showing that for large m composite permutations yield XOR tables that are optimized against at least two properties that facilitate differential attacks.

Example 1. The S-boxes of DES are mappings of the form $Si : Z_2^2 \times Z_2^4 \rightarrow Z_2^4$ *with 2 'control' bits and 4 'data' bits. For example, S3 may be written as the following* 4×16 *table, where the control bits select the row, and the data bits select the column.*

[1] In this paper the difference operator + will refer to addition in the vector space Z_2^m, though it is possible to define other difference operators.

$$S3 = \begin{bmatrix} 15 & 1 & 8 & 14 & 6 & 11 & 3 & 4 & 9 & 7 & 2 & 13 & 12 & 0 & 5 & 10 \\ 3 & 13 & 4 & 7 & 15 & 2 & 8 & 14 & 12 & 0 & 1 & 10 & 6 & 9 & 11 & 5 \\ 0 & 14 & 7 & 11 & 10 & 4 & 13 & 1 & 5 & 8 & 12 & 6 & 9 & 3 & 2 & 15 \\ 13 & 8 & 10 & 1 & 3 & 15 & 4 & 2 & 11 & 6 & 7 & 12 & 0 & 5 & 14 & 9 \end{bmatrix}.$$

□

Let $\rho : Z_2^c \times Z_2^m \rightarrow Z_2^m$ be a mapping that consists of c control bits and m data bits. The mapping ρ mapping contains 2^c m-bit permutations $\pi_i : Z_2^m \rightarrow Z_2^m$, $0 \leq i \leq 2^c - 1$, one of which is selected (or basically multiplexed) by the control bits, and a substitution is then performed on the data bits using the selected permutation[2]. Let c be bound as $1 \leq c \leq m$. Assume that for an input $X \in Z_2^{c+m}$ the first c bits (the c most significant bits) are the control bits, which we will refer to as the *control prefix* of X. Then for $X \in Z_2^{c+m}$ the expression $\left\lfloor \frac{X}{2^m} \right\rfloor$ will extract the control prefix of X. A zero control prefix is one for which $\left\lfloor \frac{X}{2^m} \right\rfloor = 0$; all other control prefixes will be called *nonzero*.

In the XOR table for ρ, let $\Lambda_\rho(\Delta X^*, \Delta Y)$ be the entry for the input difference ΔX^* and the output difference ΔY. We will show that the distribution of the random variable $\Lambda_\rho(\Delta X^*, \Delta Y)$ for the composite mapping $\rho : Z_2^{c+m} \rightarrow Z_2^m$ when ΔX^* has a nonzero control prefix takes the form

$$\Pr(\Lambda_\rho(\Delta X^*, \Delta Y) = 2k) = \sum_{\substack{p_1+p_2+\cdots+p_{2^{c-1}}=k \\ p_i \geq 0}} \prod_{i=1}^{2^{c-1}} \Pr(\lambda_i(m) = p_i). \quad (1)$$

We will prove that the $\lambda_i(m)$ are independent identically distributed (i.i.d.) random variables, described by the following probability distribution

$$\Pr(\lambda(m) = k) = \frac{1}{2^m!} \cdot \binom{2^m}{k} \cdot \frac{(2^m - k)!}{e} \cdot \left(1 + O\left(\frac{1}{(2^m - k)!}\right)\right). \quad (2)$$

This distribution is derived from the number of fixed points in an m-bit permutation (see Theorem 1). On the other hand, when ΔX^* has a zero control prefix it will be shown that

$$\Pr(\Lambda_\rho(\Delta X^*, \Delta Y) = 2k) = \sum_{\substack{p_1+p_2+\cdots+p_{2^c}=k \\ p_i \geq 0}} \prod_{i=1}^{2^c} \Pr(\Lambda_{\pi_i}(\Delta X, \Delta Y) = 2p_i) \quad (3)$$

where $\Lambda_{\pi_i}(\Delta X, \Delta Y)$ is the XOR table entry for $\Delta X, \Delta Y$ in π_i and ΔX are the m data bits of ΔX^*. Again these random variables are i.i.d. In this case, O'Connor [10, 11] has shown that $\Lambda_{\pi_i}(\Delta X, \Delta Y)$ is described by the following probability distribution

$$\Pr(\Lambda_{\pi_i}(\Delta X, \Delta Y) = 2k) = \binom{2^{m-1}}{k}^2 \cdot \frac{k! \cdot 2^k \cdot \Phi(2^{m-1} - k)}{2^m!} \quad (4)$$

[2] In this paper let ρ denote a composite permutation and let π denote a permutation.

where

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

Examining the work of Biham and Shamir [1] on DES indicates that the differential cryptanalyst is interested in two properties related to the individual XOR tables: (a) the largest entry in the XOR table; (b) the fraction of the XOR table that is zero. The value of (a) will influence the probability of the most likely characteristic, while the value (b) will influence the signal-to-noise ratio in the experiments to determine the key (see [1] for definitions and details). The system designer should then attempt to minimize the quantity in (a) and maximize the quantity in (b). Our basic result is that composite permutations are well-suited to this min-max problem.

We will model an XOR table by assuming that each entry of the table is distributed according to either eq. (1) or eq. (3), and further assume that the entries are independent. Using this model we are able to show that the number of zero entries in an XOR table for a composite permutation is well-approximated by the expression

$$(2^m - 1) \cdot \left[2 + e^{-2^{c-1}} \cdot (2^m - 1) \right] + \frac{2^{2m}(2^c - 1)}{e^{2^{c-1}}}. \tag{5}$$

By considering the cases where $\Delta X = 0$ or $\Delta Y = 0$ it is easily shown that every XOR table will have at least $2^{m+1} - 2$ entries that are zero. From eq. (5) we see that as c approaches m, the fraction of zero entries in the XOR table approaches $2^{m+1} - 2$, the least number possible (see the computational results in Table 1).

From eqs. (1) and (3) we see that $\Lambda_\rho(\Delta X, \Delta Y)$ is a sum of i.i.d. random variables. We will use the law of large numbers to show that as c is increased the expected value of $\Lambda_\rho(\Delta X, \Delta Y)$ approaches 2^c. It then follows that the probability of a characteristic for which both ΔX and ΔY are not equal to zero is approximately 2^{-m}.

2 Some notation

If a difference $X + X' \in Z_2^{c+m}$ is written as $b\Delta X$, then let b be the control prefix, ie. $\lfloor (b\Delta X)/2^m \rfloor = b$. The set of all bijective mappings $\pi : Z_2^m \to Z_2^m$ is known as the symmetric group on 2^m objects and is denoted as S_{2^m}. Let $[\cdot]$ be a boolean predicate that evaluates to 0 or 1 such as $[n$ is prime$]$.

3 Characteristics in Composite Permutations

Initially consider the case where $c = 1$ and ρ consists of two permutations π_0 and π_1, from which we will directly generalize to the cases where $c > 1$. Consider determining the pairs XOR table distribution for ρ. Let $\Delta X^* = 0\Delta X \in Z_2^{m+1}$

where $\Delta X \in Z_2^m$ such that ΔX^* has a zero control prefix. Let $\Delta X^*, \Delta Y \in Z_2^m$ be a characteristic for ρ. We then have that

$$\begin{aligned}
\Lambda_\rho(\Delta X^*, \Delta Y) &= \sum_{\substack{X, X' \in Z_2^{m+1} \\ X + X' = \Delta X^*}} [\rho(X) + \rho(X') = \Delta Y] \\
&= \sum_{\substack{X, X' \in Z_2^m \\ X + X' = \Delta X}} [\pi_0(X) + \pi_0(X') = \Delta Y] \; + \; [\pi_1(X) + \pi_1(X') = \Delta Y] \\
&= \Lambda_{\pi_0}(\Delta X, \Delta Y) + \Lambda_{\pi_1}(\Delta X, \Delta Y).
\end{aligned}$$

Then for a zero control prefix, the probability of the characteristic $\Delta X^*, \Delta Y$ will be the average of the probabilities for the characteristic $\Delta X, \Delta Y$ in π_0 and π_1. On the other hand, consider the case where $\Delta X^* = 1\Delta X \in Z_2^m$. Then we have that

$$\begin{aligned}
\Lambda_\rho(\Delta X^*, \Delta Y) &= \sum_{\substack{X, X' \in Z_2^{m+1} \\ X + X' = \Delta X^*}} [\rho(X) + \rho(X') = \Delta Y] \\
&= \sum_{\substack{X < X', \; X, X' \in Z_2^{m+1} \\ X + X' = \Delta X^*}} 2 \cdot [\rho(X) + \rho(X') = \Delta Y] \\
&= \sum_{\substack{X < X', \; X, X' \in Z_2^{m+1} \\ X + X' = \Delta X^*}} 2 \cdot [\pi_0(X) + \pi_1(X') = \Delta Y] \\
&\stackrel{\text{def}}{=} 2 \cdot \lambda(m). \qquad (6)
\end{aligned}$$

In the theorem below we prove that the expected value of $\lambda(m)$ approaches unity.

Theorem 1. Assuming that π_0 and π_1 are selected independently and uniformly from S_{2^m}

$$\mathbf{E}[\lambda(m)] = 1 + O\left(\frac{e^{2^m}}{2^{(m-1)\cdot 2^m}}\right).$$

Proof. From eq. (6) we have that

$$\begin{aligned}
\lambda(m) &= \sum_{\substack{X < X', \; X, X' \in Z_2^{m+1} \\ X + X' = \Delta X}} [\pi_0(X) + \pi_1(X') = \Delta Y] \\
&= \sum_{\substack{X < X', \; X, X' \in Z_2^{m+1} \\ X + X' = \Delta X}} \Pr(\pi_0(X) = \alpha \mid \pi_1(X') + \Delta Y = \alpha) \qquad (7)
\end{aligned}$$

since this conditional probability in eq. (7) is either 1 or 0 for all X, X'. Notice that all choices for π_1 are equivalent in that there is a unique solution to $\pi_1(X') + \Delta Y = \alpha$ for any fixed X' and ΔY. Then without loss of generality assume that π_1 is the identity permutation, $\pi_1(a) = a$, $a \in Z_2^m$. Also since π_0 and π_1 are

independent, then without loss of generality assume that $\Delta X = \Delta Y = 0$. It then follows that

$$\begin{aligned}
\lambda(m) &= \sum_{\substack{X<X',\ X,X'\in Z_2^{m+1} \\ X+X'=\Delta X}} \Pr(\pi_0(X) = \alpha \mid \pi_1(X') = \alpha) \\
&= \sum_{\substack{X<X',\ X,X'\in Z_2^{m+1} \\ X+X'=\Delta X}} \Pr(\pi_0(X) = X' \mid \pi_1(X') = X') \\
&= \sum_{\substack{X<X',\ X,X'\in Z_2^{m+1} \\ X+X'=\Delta X}} [\pi_0(X) = X'] \\
&= \sum_{X\in Z_2^m} [\pi_0(X) = X]
\end{aligned}$$

where the last two simplifications follow from the fact that $X = X'$ since $\Delta X = 0$, and π_1 is the identity permutation. It then follows that $\lambda(m)$ is equivalent to the number of fixed points $\pi_0(a) = a$ for $a \in Z_2^m$. A permutation that has no fixed points is called a *derangement*. It is well-known [7] that the number of m-bit permutations π that are derangements D_n is given as

$$D_n = 2^m! \cdot \sum_{i=0}^{2^m} \frac{(-1)^i}{i!} = \frac{2^m!}{e} \cdot \left(1 + O\left(\frac{1}{2^m!}\right)\right).$$

It then follows that for large m

$$\mathbf{E}[\lambda(m)] = \frac{1}{2^m!} \cdot \sum_{k=0}^{2^m} k \cdot \binom{2^m}{k} \cdot \frac{(2^m - k)!}{e} \cdot \left(1 + O\left(\frac{1}{(2^m - k)!}\right)\right)$$

which simplifies to

$$\begin{aligned}
\mathbf{E}[\lambda(m)] &= \frac{1}{e} \cdot \sum_{k=0}^{2^m-1} \frac{1}{k!} \cdot \left(1 + O\left(\frac{1}{(2^m - k)!}\right)\right) \\
&= \frac{1}{e} \cdot \left[\sum_{k=0}^{2^m-1} \frac{1}{k!} + \sum_{k=0}^{2^m-1} O\left(\frac{1}{(2^m - k)! \cdot k!}\right)\right] \\
&= \frac{1}{e} \cdot \left[e + O\left(\frac{1}{2^m!}\right) + O\left(\frac{2^m}{(2^{m-1}!)^2}\right)\right] \qquad (8) \\
&= 1 + O\left(\frac{e^{2^m}}{2^{(m-1)\cdot 2^m}}\right).
\end{aligned}$$

The simplification in eq. (8) follows from the fact that the summands $\sum_{k=0}^{2^m} \frac{1}{k!\cdot(2^m-k)!}$ are unimodal and symmetric. □

Corollary 3.1 $\Pr(\lambda(m) = 0) = e^{-1} + o(1)$.

Proof. From the previous theorem, the probability that $\lambda(m)$ is zero is equal to the probability that an m-bit permutation is a derangement, which is $e^{-1} + O\left(\frac{1}{2^m!}\right)$. □

We have now computed the exact distribution of an entry in the XOR table corresponding to a characteristic with a nonzero control prefix when $c = 1$. It follows that

$$\mathbf{E}[\Lambda_\rho(\Delta X^*, \Delta Y)] = 2 \cdot \mathbf{E}[\lambda(m)]$$
$$\Pr(\Lambda_\rho(\Delta X^*, \Delta Y) = 0) = e^{-1} + o(1)$$

since $\Pr(2 \cdot \lambda(m) = 0) = \Pr(\lambda(m) = 0/2) = \Pr(\lambda(m) = 0)$.

We are able to generalize our results for $c > 1$. Let ρ consist of 2^c permutations selected independently and uniformly from S_{2^m}. Let $\Delta X^* = b\Delta X$, $b \neq 0 \in Z_2^c$. By definition we have that

$$\begin{aligned}
\Lambda_\rho(\Delta X^*, \Delta Y) &= \sum_{\substack{X, X' \in Z_2^{c+m} \\ X + X' = \Delta X^*}} [\rho(X) + \rho(X') = \Delta Y] \\
&= \sum_{\substack{X, X' \in Z_2^m \\ X + X' = \Delta X}} \sum_{\substack{a + a' = b \in Z_2^c \\ b\Delta X = \Delta X^*}} [\pi_a(X) + \pi_{a'}(X') = \Delta Y] \\
&= \sum_{\substack{a + a' = b \in Z_2^c \\ b\Delta X = \Delta X^*}} \sum_{\substack{X < X' \\ X, X' \in Z_2^m \\ X + X' = \Delta X}} 2 \cdot [\pi_a(X) + \pi_{a'}(X') = \Delta Y] \\
&\stackrel{\text{def}}{=} 2 \cdot \sum_{k=1}^{2^{c-1}} \lambda_{i,j,k}(m) \qquad (9)
\end{aligned}$$

where $\Delta X^* = i$ and $\Delta Y = j$. For fixed i, j in the range $1 \leq i, j \leq 2^m - 1$ and $1 \leq k \leq 2^{c-1}$, the $\lambda_{i,j,k}(m)$ are independent and are distributed identically as $\lambda(m)$ from eq. (6). When ΔX^* has a zero control prefix it follows that

$$\begin{aligned}
\Lambda_\rho(\Delta X^*, \Delta Y) &= \sum_{\substack{X, X' \in Z_2^m \\ X + X' = \Delta X}} \sum_{a \in Z_2^c} [\pi_a(X) + \pi_a(X') = \Delta Y] \\
&= \sum_{i=0}^{2^c - 1} \Lambda_{\pi_i}(\Delta X, \Delta Y).
\end{aligned}$$

4 Joint distributions

Let the XOR table for ρ be denoted as $\Lambda_\rho(i, j)$ for $0 \leq i < 2^{c+m}$, $0 \leq j < 2^m$, where i and j are interpreted as binary strings of length $c + m$ and m

respectively. We may then define the XOR table in terms of its characteristics $\Delta X^* = i, \Delta Y = j$ as follows:

$$\Lambda_\rho(i,j) = 2 \cdot \sum_{k=1}^{2^{c-1}} \lambda_{i,j,k}(m) \qquad \left\lfloor \frac{i}{2^m} \right\rfloor > 0 \tag{10}$$

$$\Lambda_\rho(i,j) = \sum_{k=1}^{2^c} \Lambda_{\pi_k}(i,j) \qquad \left\lfloor \frac{i}{2^m} \right\rfloor = 0 \tag{11}$$

where the distribution for $\Lambda_\pi(i,j)$ is given in eq. (4) and $\sum_{k=1}^{2^{c-1}} \lambda_{i,j,k}(m)$ is defined in eq. (9). In analyzing properties of the XOR table for a composite permutation, we are concerned with the joint distribution of the $\Lambda_\rho(i,j)$, which in turn, is the joint distribution of the $\lambda_{i,j,k}(m)$. Observe that for fixed i,j and varying k the $\lambda_{i,j,k}(m)$ are *independent*, but for varying i,j,k the $\lambda_{i,j,k}(m)$ are *dependent*. However, for sake of analysis, *we will assume the* $\lambda_{i,j,k}(m)$ *to be independently distributed.* That is, we will assume that the individual XOR table entries are independently distributed. This assumption allows the probability of events for the XOR table, such as the size of the largest entry, to be cast in terms of events for the individual table entries. Results obtained by experimentation presented below show that this assumption leads to only a small deviation from the actual value of an XOR entry (see Table 1).

Theorem 2. Let the composite permutation $\rho : Z_2^{c+m} \rightarrow Z_2^m$ consist of 2^c independently and uniformly selected m-bit permutations. Let $\Lambda_{\rho,0}$ be the number of entries in the XOR table that are expected to be zero. Then

$$\Lambda_{\rho,0} \sim (2^m - 1) \cdot \left[2 + e^{-2^{c-1}} \cdot (2^m - 1) \right] + \frac{2^{2m}(2^c - 1)}{e^{2^{c-1}}}. \tag{12}$$

□

Comparisons between $\Lambda_{\rho,0}$ as derived in Theorem 2 and the observed fraction $\overline{\Lambda_{\rho,0}}$ of zero entries in the XOR table of m_p random composite permutations ρ are given in Table 1. The table indicates that the estimates $\Lambda_{\rho,0}$ from Theorem 2 are very accurate, which validates the assumption that the distribution of the $\lambda_{i,j,k}(m)$ is independent.

Biham and Shamir [1] report that 20%–30% of the entries in the S-boxes of DES are zero. Theorem 2 yields that approximately 16% of the XOR table entries in a mapping with 4 data bits and 2 control bits are expected to be zero; further, a random sample of 10,000 such mappings has yielded an average of 15.7% zero entries in the corresponding XOR tables. This suggests that the set of design criteria for the S-boxes has increased the expected density of zeroes in the XOR table.

The (strong) law of large number states that for random variables $Y_1, Y_2, \ldots, Y_N$ which are i.i.d. with mean $\mathbf{E}[Y]$

$$\left| \frac{Y_1 + Y_2 + \cdots + Y_N}{N} - \mathbf{E}[Y] \right| < \epsilon \tag{13}$$

m	c	$\Lambda_{\rho,0}/2^{c+2m}$	$\overline{\Lambda_{\rho,0}}$	m_p
4	1	0.40419	0.39433	10000
4	2	0.16053	0.15707	10000
4	3	0.03268	0.03221	10000
4	4	0.00765	0.00765	10000
5	1	0.38683	0.38160	1000
5	2	0.14839	0.14666	1000
5	3	0.02574	0.02544	1000
5	4	0.00411	0.00410	1000
5	5	0.00189	0.00189	1000
6	1	0.37755	0.37486	1000
6	2	0.14197	0.14106	1000
6	3	0.02208	0.02196	1000
6	4	0.00225	0.00225	1000

Table 1. Estimates of $\Lambda_{\rho,0}$ for composite permutations.

for any $\epsilon > 0$ as N becomes large. That is, the sample mean approaches the expectation of the random variable Y_i. Observe that an individual entry of the XOR table for a multiple permutation is a sum of i.i.d. random variables. If $\Delta X^* = i, \Delta Y = j$ then for characteristics with zero control prefix, the XOR entry is a sum of 2^c random variables $\Lambda_{\pi_k}(i,j)$ as defined in eq. (11), and when the control prefix is nonzero, the XOR entry is a sum of 2^{c-1} random variables $\lambda_{i,j,k}(m)$ as defined in eq. (10). The mean of $\lambda_{i,j,k}(m)$ was proven to be $1+o(1)$ in Theorem 1. It can also be shown [11] from eq. (4) that the mean of $\Lambda_{\pi_k}(i,j)$ is $1+o(1)$. Both these $o(1)$ terms dominate $2^c \leq 2^m$ and the largest value in the table for large m will be $2^c + o(1)$ given our independence assumption.

5 Conclusion

Our analysis has shown that for sufficiently large c, a very small fraction of the XOR table will be zero, and that the largest entry will be close to $2^c + o(1)$. We note that no special algorithms are required to construct composite permutation ρ that have these properties as they are a consequence of the law of large numbers. For this reason it appears that composite permutations provide are more resistant to differential attacks than most other mappings, including single permutations π.

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 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.
3. 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.
4. 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.
5. M. H. Dawson and S. E. Tavares. An expanded set of S-box design criteria based on information theory and its relation to differential-like attacks. *Advances in Cryptology, EUROCRYPT 91, Lecture Notes in Computer Science, vol. 547, D. W. Davies ed., Springer-Verlag*, pages 352–367, 1991.
6. W. Feller. *An Introduction to Probability Theory with Applications.* John Wiley and Sons, 3rd edition, Volume 1, 1968.
7. R. L. Graham, D. E. Knuth, and O. Patshnik. *Concrete Mathematics, A Foundation for Computer Science.* Addison Wesley, 1989.
8. 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.
9. 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.
10. L. J. O'Connor. On the distribution of characteristics in bijective mappings. presented at Eurocrypt 93, Norway, May 1993. Also accepted for publication in the *Journal of Cryptology.*
11. L. J. O'Connor. *An analysis of product ciphers using boolean functions.* PhD thesis, Department of Computer Science, University of Waterloo, 1992.
12. 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.
13. A. Sorkin. LUCIFER: a cryptographic algorithm. *Cryptologia*, 8(1):22–35, 1984.