Linear Approximation of Block Ciphers

Kaisa Nyberg
Prinz Eugen-Straße 18/6, A-1040 Vienna, Austria

May 27, 1994

Abstract

The results of this paper give the theoretical fundaments on which Matsui's linear cryptanalysis of the DES is based. As a result we obtain precise information on the assumptions explicitely or implicitely stated in [2] and show that the success of Algorithm 2 is underestimated in [2]. We also derive a formula for the strength of Algorithm 2 for DES-like ciphers and see what is its dependence on the plaintext distribution. Finally, it is shown how to achieve proven resistance against linear cryptanalysis.

1 Linear Cryptanalysis of a DES-like Cipher

We consider a DES-like iterated cipher consisting of r rounds of iteration

$$\begin{aligned} X_L(i+1) &= X_R(i) \\ X_R(i+1) &= X_L + f(E(X_R(i)) + K_i) \end{aligned}$$

at the rounds $i = 1, 2, \ldots, r-1$, and

$$\begin{aligned} C_L &= X_L(r) + f(E(X_R(r)) + K_r) \\ C_R &= X_R(r) \end{aligned}$$

Here we have denoted by K_i the round key used at the ith round and by $X(i) = (X_L(i), X_R(i))$ the input to the ith round with its left and right halves. Hence $X(1) = P = (P_L, P_R)$ is the plaintext and $C = (C_L, C_R)$ is the ciphertext.
In [2] M. Matsui introduces the linear cryptanalysis method to recover with high probability certain key bits using sufficient large number of known plaintext-ciphertext pairs. The main part of this attack is a procedure called Algorithm 2 which can be used to recover 12 bits of a DES-key. Let us give a short description of this procedure.
First the round function is analyzed to find linear approximations of the function f of the form

$$b(i) \cdot f(Z) = c_i \cdot Z$$

which holds with probability p_i over the uniform distribution of the random variable Z and such that $|\, p_i - \frac{1}{2} \,|$ is non-negligible. Then $r-2$ of such approximations are chained to obtain a linear approximation over $r-2$ rounds from the second to the second last round of the form

$$a \cdot X + b \cdot Y + c \cdot (k_2, \ldots, k_{r-1}) \;=\; 0 \tag{1}$$

where $X = X(2)$ and $Y = X(r)$ and $k = (k_2, \ldots, k_{r-1})$ is the vector formed by concatenating the unknown round keys k_i used at the rounds $i = 2, \ldots, r-1$. The probability of (1) over the distribution of X is denoted by $p(a, b, c; k)$. This probability should not be equal to $\frac{1}{2}$. In his analysis Matsui implicitely assumes that that the inputs to f at different rounds are independent and uniformly random and obtains an estimate

$$p(a, b, c; k) \;\approx\; \frac{1}{2} + 2^{r-3} \prod_{i=2}^{r-1} (p_i - \frac{1}{2}) \tag{2}$$

using the classical "piling-up" lemma. Let us denote by $p(a, b, c)$ the average of $p(a, b, c; k)$ taken over k. If the round keys K_i are independent and uniformly random then the inputs to f at each round are independent and uniformly random and the right hand side of (2) equals to the probability of the linear approximate relation

$$a \cdot X + b \cdot Y + c \cdot (K_2, \ldots, K_{r-1}) = 0. \tag{3}$$

But the probability of (3) is the average probability of (1). Hence by (2) it is essentially estimated that

$$p(a, b, c; k) \;\approx\; p(a, b, c) \tag{4}$$

for almost all k.
The next step in Algorithm 2 is to substitute in (1)

$$\begin{aligned} X &= (P_R, P_L + f(E(P_R) + k_1) \\ Y &= (C_L + f(E(C_R) + k_r), C_R) \end{aligned}$$

to achieve the following approximate relation

$$\begin{aligned} & a_L \cdot P_R + a_R \cdot P_L + b_L \cdot C_L + b_R \cdot C_R \\ + \;& a_R \cdot f(E(P_R) + k_1) + b_L \cdot f(E(C_R) + k_r) \\ + \;& c \cdot (k_2, \ldots, k_{r-1}) = 0 \end{aligned} \tag{5}$$

which holds with probability $p(a, b, c; k)$ if k_1 and k_r are the correct round keys at the first and the last rounds. But if either k_1 or k_r is incorrect then it is hypothetized that the uncertainty of (5) increases. In DES the function f constitutes of eight parallel substitutions with six bit inputs each. Therefore it is possible to design (1) in such a way that only six bits of k_1 and six bits of k_r are involved in (5). For each possible 12-bit combination the cryptanalyst, who is given N different known plaintext-ciphertext pairs, counts the number N_0 of plaintexts for which

$$\begin{aligned} & a_L \cdot P_R + a_R \cdot P_L + b_L \cdot C_L + b_R \cdot C_R \\ + \quad & a_R \cdot f(E(P_R) + k_1) + b_L \cdot f(E(C_R) + k_r) = 0 \end{aligned}$$

holds. The 12-bit candidate is accepted that maximizes the quantity

$$\left| \frac{N_0}{N} - \frac{1}{2} \right|$$

Note that this step is independent of the vector c in (5) which selects certain bits from the round keys $k_2, \ldots, k_{r-1}$.
In [2] Matsui shows that in order to achieve a predetermined success rate for Algorithm 2 the number N of known plaintext needed in the cryptanalysis is inversely proportional to $| p(a,b,c;k) - \frac{1}{2} |^2$. Based on the estimate (4) Matsui obtains

$$| p(a,b,c;k) - \frac{1}{2} | \quad \approx \quad | p(a,b,c) - \frac{1}{2} | \tag{6}$$

for practically all k and the chosen value of c. The main purpose of this work is to show that (6) and (4) do not hold in general. The Fundamental Theorem to be proved in Section 2 implies that the average of $| p(a,b,c;k) - \frac{1}{2} |^2$ over k equals to the sum of $| p(a,b,c) - \frac{1}{2} |^2$ over c. This sum is in general strictly larger than $| p(a,b,c) - \frac{1}{2} |^2$ for any c. These values could be equal only in the case when there is only one c, i.e., one chain of round approximations, which gives a non-negligible positive value of $| p(a,b,c) - \frac{1}{2} |$.
It follows that the average success rate of Algorithm 2 is larger than estimated by Matsui in [2]. On the other hand, the success of Matsui's Algorithm 1 essentially depends on the assumption (4) and may be significantly weakened if there are more than one c with non-negligible value $| p(a,b,c) - \frac{1}{2} |$.
We conclude that Algorithm 2 makes in fact use of a family of linear approximate expressions

$$a \cdot X + b \cdot Y + c \cdot (K_2, \ldots, K_{r-1})$$

where a and b are fixed but c varies. This means that the round approximations, which uniquely determine c and are uniquely determined by c, can be chosen in all possible ways to form a chain of approximations from $a \cdot X$ to $b \cdot Y$. Hence there is a close analog with what is called differentials in differential cryptanalysis [1].
In Section 2 we discuss the theory of linear approximation of block ciphers and prove a version of Parseval's theorem. Based on this theorem we give a definition of approximate linear hull of a block cipher and its potential. In Section 3 we determine the potential of the approximate linear hull for DES-like ciphers in terms of the probabilities of the approximations of the function f at each round. Finally, in Section 4 we show that with highly nonlinear f one can achieve proven resistance against linear cryptanalysis attack. The proofs of the results presented in Sections 3 and 4 are omitted due to space constraints.

2 Linear Approximation of a Function of Two Random Variables

Let $\mathsf{F} = GF(2)$ be the finite field of order two. Let $X \in \mathsf{F}^m$ and $K \in \mathsf{F}^\ell$ be random variables and $Y = Y(X, K)$, $Y \in \mathsf{F}^n$, be a random variable which is a function of X and K. Then we have the following generalisation of Parseval's theorem.

Theorem 1 (The Fundamental Theorem) *If X and K are independent and K is uniformly distributed, then for all $a \in \mathsf{F}^m$, $b \in \mathsf{F}^n$ and $\gamma \in \mathsf{F}^\ell$*

$$2^{-\ell} \sum_{k \in \mathsf{F}^\ell} | P_X(a \cdot X + b \cdot Y(X;k) = 0) - \frac{1}{2} |^2 =$$

$$2^{-\ell} \sum_{k \in \mathsf{F}^\ell} | P_X(a \cdot X + b \cdot Y(X;k) + \gamma \cdot k = 0) - \frac{1}{2} |^2 =$$

$$\sum_{c \in \mathsf{F}^\ell} | P_{X,K}(a \cdot X + b \cdot Y(X;K) + c \cdot K = 0) - \frac{1}{2} |^2$$

Proof. Since this theorem holds without the assumption of the independence of X and K we give the proof in the general case.
Let us first recall that for a Boolean function g of n binary variables and for a random variable $Z \in \mathsf{F}^n$ we have

$$\sum_z P_Z(Z = z)(-1)^{g(z)} \quad = \quad 2P_Z(g(Z) = 0) - 1.$$

Applying this simple equality first to the random variable $Z = (X, K)$ and then to the random variable $Z = (X|K)$, we obtain

$$\begin{aligned}
& \sum_{c \in \mathsf{F}^\ell} | P_{X,K}(a \cdot X + b \cdot Y(X,K) + c \cdot K = 0) - \frac{1}{2} |^2 \\
= \; & \frac{1}{4} \sum_{c \in \mathsf{F}^\ell} (\sum_{k \in \mathsf{F}^\ell} \sum_{x \in \mathsf{F}^m} P_{X,K}(X = x, K = k)(-1)^{a \cdot x + b \cdot y(x,k) + c \cdot k})^2 \\
= \; & \frac{1}{4} \sum_{c \in \mathsf{F}^\ell} \sum_{k,\gamma \in \mathsf{F}^\ell} \sum_{x,\xi \in \mathsf{F}^m} P_{X,K}(X = x, K = k)(-1)^{a \cdot x + b \cdot y(x,k) + c \cdot k} \\
& \cdot P_{X,K}(X = \xi, K = \gamma)(-1)^{a \cdot \xi + b \cdot y(\xi,\gamma) + c \cdot \gamma} \\
= \; & 2^{-2\ell-2} \sum_{k,\gamma \in \mathsf{F}^\ell} \sum_{x,\xi \in \mathsf{F}^m} P_{X,K}(X = x | K = k)(-1)^{a \cdot x + b \cdot y(x,k)} \\
& \cdot P_{X,K}(X = \xi | K = \gamma)(-1)^{a \cdot \xi + b \cdot y(\xi,\gamma)} \sum_{c \in \mathsf{F}^\ell} (-1)^{c \cdot (k + \gamma)} \\
= \; & 2^{-\ell-2} \sum_{k \in \mathsf{F}^\ell} (\sum_{x \in \mathsf{F}^m} P_{X,K}(X = x | K = k)(-1)^{a \cdot x + b \cdot y(x,k)})^2
\end{aligned}$$

$$= 2^{-\ell} \sum_{k \in \mathsf{F}^\ell} | P_{X,K}(a \cdot X + b \cdot Y(X,K) = 0 \,|\, K = k) - \frac{1}{2} |^2$$

since

$$\sum_{c \in \mathsf{F}^\ell} (-1)^{c \cdot (k+\gamma)} = \begin{cases} 0 & \text{for } k \neq \gamma \\ 2^\ell & \text{for } k = \gamma. \end{cases}$$

In our application $Y = Y(X, K)$ is a block cipher, or some rounds of it, and X is the plaintext and K the uniformly distributed key. We assume, as usual, that the plaintex and the key are independent. Let us introduce the following notation:

$$\begin{aligned} pot(a,b\,;k) &= | P_X(a \cdot X + b \cdot Y(X,k) + c \cdot k = 0) - \frac{1}{2} |^2 \\ &= | P_X(a \cdot X + b \cdot Y(X,k) = 0) - \frac{1}{2} |^2 \\ pot(a,b,c) &= | P_{X,K}(a \cdot X + b \cdot Y(X,K) + c \cdot K = 0) - \frac{1}{2} |^2 \end{aligned}$$

The quantity $pot(a, b\,; k)$ is called the *potential of the linear approximate expression* $a \cdot X + b \cdot Y(X, k)$ *for key* k. The quantity $pot(a, b, c)$ is called the *potential of the linear approximate expression* $a \cdot X + b \cdot Y + c \cdot K$. Further we can interpret the sum of $pot(a, b, c)$ over c as the potential of the family of linear approximate expressions

$$a \cdot X + b \cdot Y + c \cdot K, \quad c \in \mathsf{F}^\ell$$

We call this family the approximate linear hull $ALH(a, b)$ of the block cipher $Y = Y(X, K)$ determined by a and b. Using this terminology we can express the result of the Fundamental Theorem as follows: the average potential of the linear approximate expression $a \cdot X + b \cdot Y(X, k)$ over the keys is the potential of the corresponding approximate linear hull $ALH(a, b)$ of the cipher $Y = Y(X, K)$.

3 Linear Approximation of a DES-like Cipher

In this section we represent the potential of $ALH(a, b)$ of a DES-like cipher in the terms of the probabilities of the round approximations. We make use of the notation introduced in Section 1 and assume that f is a function from F^m to F^n, $m \geq n$, and the expansion mapping E from F^n to F^m is linear. Let E^t be the transpose of E. We have the following

Theorem 2 *If the round keys of r rounds of a DES-like cipher are independent and uniformly random then $\ell = mr$ and for all a and b the potential of $ALH(a, b)$ equals*

$$4^r \sum_{c \in \mathsf{F}^\ell} | P_X((a + b^0) \cdot X = 0) - \frac{1}{2} |^2 \prod_{i=1}^{r} | P_Z(b_R^i \cdot f(Z) = c_i \cdot Z) - \frac{1}{2} |^2$$

where

$$b^r = (b_L, b_R),\ \ b^{i-1} = (b_R^i, b_L^i + E^t(c_i)),\ \ \textit{for } i = 1, 2, \ldots, r, \textit{ and}$$
$$c = (c_1, \ldots, c_r).$$

This representation of the potential of an $ALH(a, b)$ shows the role of the plaintext distribution. Particularly, if the plaintext is uniformly random then the summation can be taken over all $c \in \mathsf{F}^\ell$ such that

$$a_L + b_L + \sum_{i=1}^{\frac{r}{2}} E^t(c_{2i}) = 0 \text{ and } a_R + b_R + \sum_{i=1}^{\frac{r-1}{2}} E^t(c_{2i-1}) = 0$$

(assuming that r is even), since for all other c we have $pot(a, b, c) = 0$. If c satisfies these equations we denote $c \in S(a, b)$. In this case the potential of $ALH(a, b)$ equals

$$4^{r-1} \sum_{c \in S(a,b)} \prod_{i=1}^{r} | P_Z(b_R^i \cdot f(Z) = c_i \cdot Z) - \frac{1}{2} |^2$$

4 Resistance Against Linear Cryptanalysis

The *linearity* of a function $f : \mathsf{F}^m \rightarrow \mathsf{F}^n$ is defined as

$$\mathcal{L}(f) = 2 \max_{a \text{ any}, b \neq 0} | P_Z(b \cdot f(Z) = a \cdot Z) - \frac{1}{2} | = 1 - 2^{1-m} \mathcal{N}(f)$$

where Z is uniformly random in F^m and $\mathcal{N}(f)$ is the nonlinearity of f (see e.g. [3]). Based on Theorem 2 we get the following

Theorem 3 *For r rounds, $r \geq 4$, of a DES-like cipher with independent round keys and uniformly random plaintext*

$$2^{-\ell} \sum_{k \in \mathsf{F}^\ell} | P_X(a \cdot X + b \cdot Y(X; k) = 0) - \frac{1}{2} |^2 \leq 2^{2(m-n)-1} \mathcal{L}(f)^4$$

Examples of functions of f which give proven resistance against both differential and linear cryptanalysis can be found e.g. in [3].

Acknowledgement. I would like to thank Lars Knudsen for numerous discussions on linear cryptanalysis and for proposing to look for counterparts of differentials in the linear cryptanalysis method.

References

[1] X. Lai, J. L. Massey, S. Murphy, ***Markov ciphers and differential cryptanalysis***, Advances in Cryptology – EUROCRYPT'91, Lecture Notes in Computer Science **547**, Springer-Verlag, 1992.

[2] M. Matsui, ***Linear cryptanalysis method for DES cipher***, in Advances in Cryptology – EUROCRYPT'93, Lecture Notes in Computer Science **765**, Springer-Verlag, 1994, pp. 386-397.

[3] K. Nyberg, ***Differentially uniform mappings for cryptography***, ibidem, pp. 55-64