On Unbiased Linear Approximations

Jonathan Etrog* and Matthew J.B. Robshaw

Orange Labs
38–40 rue du Général Leclerc
92794 Issy les Moulineaux Cedex 9, France
forename.surname@orange-ftgroup.com

Abstract. In this paper we explore the recovery of key information from a block cipher when using unbiased linear approximations of a certain form. In particular we develop a theoretical framework for their treatment and we confirm their behaviour with experiments on reduced-round variants of DES. As an application we show a novel form of linear cryptanalysis using multiple linear approximations which can be used to extract key information when all pre-existing techniques would fail.

Keywords: linear cryptanalysis, multiple approximations, entangled approximations.

1 Introduction

The technique of linear cryptanalysis [12,13] has become a standard tool in symmetric cryptanalysis. When used with block ciphers the essential idea is to find a linear expression or *linear approximation* that links bits of the plaintext, the ciphertext, and the key and which holds with some probability $\frac{1}{2} + s = \frac{1}{2}(1 + \epsilon)$, where the value s is known as the *bias* and $\epsilon = 2s$ is known as the *imbalance* or *correlation* [5]. If the bias or imbalance of the linear approximation is zero, that is $s = \epsilon = 0$, then we say that the linear approximation is *balanced* or *unbiased*.

As is well-known, a linear approximation can be used to recover one bit of key information if the bias or imbalance of the linear approximation is nonzero, that is $s, \epsilon \neq 0$. While elements of doing this were first described in [18], Matsui's *Algorithm 1* [12] describes how to recover this bit of key information given sufficiently many known plaintext-ciphertext pairs. The more complex *Algorithm 2* [12] uses a linear approximation as a reduced-round distinguisher and allows the recovery of more bits of key from more rounds of the cipher, and its effectiveness has been considered in [3,9,16].

The simultaneous use of *multiple linear approximations* to find a single bit of key information was first proposed in [10]. Such an approach can lead to a reduction in the data-complexity of the attack when compared to the use of a single linear approximation. Further papers exploring and extending the use of multiple linear approximations, under different assumptions and different types of approximations, include [1,4,6,7,8,11].

* Partially supported by the national research project RFIDAP ANR-08-SESU-009-03.

R. Steinfeld and P. Hawkes (Eds.): ACISP 2010, LNCS 6168, pp. 74–86, 2010.

As early as 1994, an interesting experimental phenomenon in the use of multiple linear approximations was described in [11]. There it was observed that four linear approximations could still give an advantage over using three approximations even though the fourth approximation was the linear algebraic sum of the other three. Whilst this issue has been revisited theoretically and experimentally, for instance [1,4], Murphy stressed the importance of considering the dependencies between linear approximations in recent work [14]. This has been confirmed by the framework of Hermelin *et al* using what have been called *multidimensional* techniques [2,6].

In this paper we highlight a striking consequence of this work and we explore the role of unbiased but dependent linear approximations. As shorthand, we will be referring to these approximations as *entangled* approximations (their definition will be given below) and, using [14], we will illustrate how to derive key information from such linear approximations. We will give a full theoretical treatment of this analysis, including error rates and the data requirements for key recovery, and we will also present the results of confirming experiments that have been carried out on reduced-round DES (which is a normal target cipher for experiments in the field).

We stress that this new technique is not intended as a universal replacement for traditional linear cryptanalysis, using either single or multiple approximations. Indeed techniques using biased approximations, when applicable, are likely to offer the best avenue for attack. However the results in this paper are important for two reasons:

1. We show that it is possible to use **only** the counts related to unbiased linear approximations in a linear cryptanalytic attack and still recover key information.
2. We demonstrate a practical situation where biased approximations cannot be used and all current linear cryptanalytic techniques would fail. Nevertheless, the presence of unbiased linear approximations still compromises the cipher.

2 Linear Approximations

Throughout the paper we consider a block cipher that operates on b-bit blocks, and we suppose that all the subkeys for the block cipher are concatenated to give an expanded-key $\mathbf{k}$ for the block cipher. For plaintext $\mathbf{p}$ and ciphertext $\mathbf{c}$, a linear approximation holding with probability $\frac{1}{2}(1+\epsilon)$ is usually written in the following way

$$\gamma_{\mathbf{p}} \cdot \mathbf{p} \oplus \gamma_{\mathbf{c}} \cdot \mathbf{c} = \gamma_{\mathbf{k}} \cdot \mathbf{k},$$

where $\gamma_{\mathbf{p}}$, $\gamma_{\mathbf{c}}$, and $\gamma_{\mathbf{k}}$ represent bit-masks for the plaintext, ciphertext, and expanded key. For convenience, we rewrite this linear expression by considering the plaintext and ciphertext masks as a combined single $2b$-bit data mask α, and we use $\gamma = \gamma_k$ to give

$$\alpha^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = \gamma^T \mathbf{k}.$$

In its most basic form [10], an analysis using multiple linear approximations uses a collection $\alpha_1, \ldots, \alpha_l$ of data (plaintext-ciphertext) masks relating to the same[1] bit $k = \gamma_{\mathbf{k}}^T \mathbf{k}$ of key information to obtain the collection of approximations

$$\begin{aligned} \alpha_1^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} &= k \text{ with probability } \frac{1}{2}(1+\epsilon_1), \\ &\vdots \\ \alpha_l^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} &= k \text{ with probability } \frac{1}{2}(1+\epsilon_l). \end{aligned}$$

If we have N plaintext-ciphertext pairs $\begin{pmatrix} \mathbf{p}_1 \\ \mathbf{c}_1 \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{p}_N \\ \mathbf{c}_N \end{pmatrix}$, then we can estimate the key bit k in the following way. We let $V_1, \ldots, V_l$ be the counts corresponding to the data masks $\alpha_1, \ldots, \alpha_l$, that is

$$\begin{aligned} V_1 &= \# \left\{ \begin{pmatrix} \mathbf{p}_i \\ \mathbf{c}_i \end{pmatrix} \middle| \alpha_1^T \begin{pmatrix} \mathbf{p}_i \\ \mathbf{c}_i \end{pmatrix} = 0 \right\}, \\ &\vdots \\ V_l &= \# \left\{ \begin{pmatrix} \mathbf{p}_i \\ \mathbf{c}_i \end{pmatrix} \middle| \alpha_l^T \begin{pmatrix} \mathbf{p}_i \\ \mathbf{c}_i \end{pmatrix} = 0 \right\}, \end{aligned}$$

and we let $Y_1 = V_1 - \frac{N}{2}, \ldots, Y_l = V_l - \frac{N}{2}$ denote the centred counts.

For ease of exposition, we now restrict ourselves to the case of two linear approximations given by data masks α_{10} and α_{01} for which we have access to two centred counters Y_{10} and Y_{01}. Next we assume that the two approximations given by α_{10} and α_{01} are unbiased and yet the approximation given by the data mask $\alpha_{11} = \alpha_{01} + \alpha_{10}$ is biased. Further we assume that for whatever structural or operational reason, we don't have access to the counter Y_{11} that is related to α_{11} (see Section 5 for an example of this).

Considered individually, neither of the two centred counts Y_{10} and Y_{01} of Section 2 can ever give any information about the key bit k. This is reflected in the typical approaches to multiple linear approximations such as those described in [1,10] which would be unable to recover the key bit with any advantage. Furthermore, even multidimensional attacks using the unbiased span approximations α_{10} and α_{01} [2] would not be able to handle this situation since our problem statement explicitly rules out using counts based on α_{11}.

Nevertheless, the purpose of this paper is to highlight the fact that taken together, the pair of centred counts (Y_{10}, Y_{01}) can be used to recover key information. The main result we exploit in using unbiased linear approximations is Theorem 1 of [14]. This essentially states that if the linear approximation corresponding to α_{11} is biased, then the two centred counts Y_{10} and Y_{01} are correlated. We can use this correlation between the two centred counts Y_{10} and Y_{01} for unbiased linear approximations to recover key information. In particular,

[1] This condition was first relaxed in [11].

if Y_{10} and Y_{01} have the same sign, then this indicates one particular value for the key bit k is more likely; whereas if Y_{10} and Y_{01} have opposite signs, then this indicates that the other value for the key bit k is more likely.

We now formalise this notion of two unbiased linear approximations to the same key bit using notation from [14].

Definition 1. Suppose α_{10} and α_{01} are the data masks for unbiased linear approximations for the key bit k, that is

$$\alpha_{10}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = k \quad \text{and} \quad \alpha_{01}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = k$$

each with probability $\frac{1}{2}$. If $\alpha_{11} = \alpha_{10} + \alpha_{01}$ is the data mask for a biased linear approximation for key bit k, that is

$$\alpha_{11}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = (\alpha_{10} + \alpha_{01})^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = k$$

with probability $\frac{1}{2}(1 + \epsilon)$ for some $\epsilon \neq 0$, then the two unbiased linear approximations for the key bit k based on the data masks α_{10} and α_{01} are *entangled* linear approximations[2] with *entanglement* ϵ.

The goal of our paper. Now our framework has been set we can state the result of the paper. When we can **only** examine the two entangled approximations, we cannot apply typical multiple approximation techniques [1,10] due to the zero capacity of the available approximations. Further we cannot apply multidimensional techniques [6,7] since we don't have access to the full set of counts related to the full set of approximations generated by the base approximations. Nevertheless, using results in this paper, we are still able to recover key information.

3 Key Recovery Using Entangled Approximations

We assume that we have N plaintext-ciphertext pairs for a pair of entangled unbiased linear approximations with entanglement ϵ. We now derive an asymptotically optimal process for recovering the key bit k from the centred counts Y_{10} and Y_{01}. We can rewrite the entanglement condition of Definition 1 as

$$\alpha_{11}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = (\alpha_{10} + \alpha_{01})^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = 0 \text{ with probability } \frac{1}{2}\left(1 + (-1)^k \epsilon\right).$$

Section II of [14] gives a bivariate normal distribution for the asymptotic joint distribution of the normalised centred counts

$$Z_{10} = \frac{2}{\sqrt{N}} Y_{10} \quad \text{and} \quad Z_{01} = \frac{2}{\sqrt{N}} Y_{01}$$

[2] Note that entanglement is not exactly coincident with statistical dependence since we specify that the base approximations are unbiased.

using central limit theory ideas, so we have

$$\mathbf{Z} = \begin{pmatrix} Z_{10} \\ Z_{01} \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 & (-1)^k\epsilon \\ (-1)^k\epsilon & 1 \end{pmatrix} \right) .$$

Thus we can write $\mathbf{Z} \sim \mathrm{N}(0; \Sigma_k)$, where

$$\Sigma_k = \begin{pmatrix} 1 & (-1)^k\epsilon \\ (-1)^k\epsilon & 1 \end{pmatrix}, \text{ so}$$
$$\Sigma_k^{-1} = \frac{1}{1-\epsilon^2} \begin{pmatrix} 1 & -(-1)^k\epsilon \\ -(-1)^k\epsilon & 1 \end{pmatrix}.$$

We can now give the likelihood function of the key bit value k given data $\mathbf{z} = \begin{pmatrix} z_{10} \\ z_{01} \end{pmatrix}$ for the normalised centred counts as

$$\begin{aligned} L(k; z) &= (2\pi)^{-1} \left|\Sigma_k\right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{pmatrix} z_{10} & z_{01} \end{pmatrix} \Sigma_k^{-1} \begin{pmatrix} z_{10} \\ z_{01} \end{pmatrix} \right) \\ &= (2\pi)^{-1} \left(1-\epsilon^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{1}{(1-\epsilon^2)} \left(z_{10}^2 + z_{01}^2 - 2(-1)^k \epsilon z_{10} z_{01} \right) \right). \end{aligned}$$

This means that the likelihood ratio for key bit $k = 1$ versus $k = 0$ is given by

$$\begin{aligned} \Lambda(\mathbf{z}) = \frac{L(1; \mathbf{z})}{L(0; \mathbf{z})} &= \frac{\exp\left(-\frac{1}{2} \frac{1}{(1-\epsilon^2)} \left(z_{10}^2 + z_{01}^2 + 2\epsilon z_{10} z_{01} \right) \right)}{\exp\left(-\frac{1}{2} \frac{1}{(1-\epsilon^2)} \left(z_{10}^2 + z_{01}^2 - 2\epsilon z_{10} z_{01} \right) \right)} \\ &= \exp\left(-\frac{2\epsilon}{1-\epsilon^2} \, z_{10} z_{01} \right), \end{aligned}$$

so we have the following log-likelihood ratio statistic

$$\log \Lambda(z) = -\frac{2\epsilon}{1-\epsilon^2} \, z_{10} z_{01}.$$

Thus the log-likelihood ratio statistic is proportional to $P(\mathbf{z}) = z_{10} z_{01}$, or equivalently to $P(\mathbf{y}) = y_{10} y_{01}$ in terms of the unnormalised centred counts. The Neyman-Pearson Lemma [17] therefore shows that the asymptotically optimal test of $k = 0$ versus $k = 1$ is given by the sign of $P(\mathbf{z})$ or $P(\mathbf{y})$. For example, for positive entanglement ($\epsilon > 0$), we choose $k = 0$ if the centred counts Y_{10} and Y_{01} have the same sign and we choose $k = 1$ if the centred counts Y_{10} and Y_{01} have the opposite signs, and we swap these choices for negative entanglement.

3.1 Success Rates in Using Entangled Approximations

We now calculate the success rates in using the process of Section 3 to recover a key bit using a pair of entangled unbiased linear approximations. Without loss of generality, we suppose that the true value of the key bit k is 0 and that the

entanglement $\epsilon > 0$. The accuracy probability $a(\epsilon)$ that this process correctly identifies the key bit value k as 0 is then given by $a(\epsilon) = \mathbf{P}(P(\mathbf{z}) > 0) = \mathbf{P}(z_{10}z_{01} > 0) = 2\mathbf{P}(z_{10}, z_{01} > 0)$. Thus this accuracy probability is given by

$$a(\epsilon) = 2\int_0^\infty \int_0^\infty \frac{1}{2\pi(1-\epsilon^2)^{\frac{1}{2}}} \exp\left(-\left(\frac{z_{10}^2 + z_{01}^2 - 2\epsilon z_{10}z_{01}}{2(1-\epsilon^2)}\right)\right) dz_{10}dz_{01},$$

where the integrand is the joint density function of normalised centred count vector $\mathbf{Z}$. The accuracy and error probabilities for certain entanglements ϵ are given in Table 1.

Table 1. Theoretical accuracy and error rates for single key bit recovery using entangled approximations with entanglement ϵ

Entanglement (ϵ)	0.002	0.004	0.01	0.02	0.04	0.1	0.2	0.4
Accuracy Rate	0.5006	0.5013	0.503	0.506	0.513	0.532	0.564	0.631
Error Rate	0.4994	0.4987	0.497	0.494	0.487	0.468	0.436	0.369

We are usually interested in the case where the entanglement ϵ is small. We can write $a(\epsilon) = a(0) + a'(0)\epsilon + \frac{1}{2}a''(0)\epsilon^2 + o(\epsilon^3)$ and calculate:

$$a(0) = 2\int_0^\infty \int_0^\infty \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(z_{10}^2 + z_{01}^2\right)\right) dz_{10}dz_{01} = \frac{1}{2},$$

$$a'(0) = 2\int_0^\infty \int_0^\infty \frac{1}{2\pi} z_{10}z_{01} \exp\left(-\frac{1}{2}\left(z_{10}^2 + z_{01}^2\right)\right) dz_{10}dz_{01} = \frac{1}{\pi},$$

$$a''(0) = 2\int_0^\infty \int_0^\infty \frac{1}{2\pi}(z_{10}^2 - 1)(z_{01}^2 - 1) \exp\left(-\frac{1}{2}\left(z_{10}^2 + z_{01}^2\right)\right) dz_{10}dz_{01} = 0.$$

For small ϵ, we can therefore express the accuracy probability as

$$a(\epsilon) = \tfrac{1}{2} + \tfrac{\epsilon}{\pi} + o(\epsilon^3),$$

which gives the error probability as $1 - a(\epsilon) \approx \frac{1}{2} - \frac{\epsilon}{\pi}$.

These success rates for the recovery of a single key bit k from a pair of entangled unbiased linear approximations do not explicitly depend on the number N of plaintext-ciphertext pairs. Thus the expression for success rate $a(\epsilon) = \frac{1}{2} + \frac{\epsilon}{\pi}$ is valid for any sample size N large enough for the central limit normal approximation to be valid. It would therefore be reasonable to approximate the distribution of the data class count vector as a multivariate normal random variable with, for example, $N = 64$ plaintext-ciphertext pairs. Furthermore, increasing the number of plaintext-ciphertext pairs beyond a point at which the normal approximation is reasonable does not materially improve the accuracy of the process.

3.2 Experimental Confirmation

To illustrate our analysis we consider some experiments using reduced-round versions of the DES [15], using the notation given by [12]. We first consider a reduced-round version of DES with three rounds and the following pair of linear approximations:

$$P_H[7,18,24,29] \oplus C_H[7,18,24,29] = K_1[22] \oplus K_3[22], \text{ and}$$
$$P_L[15] \oplus C_L[15] = K_1[22] \oplus K_3[22].$$

In the notation of Section 3, we have

$$\alpha_{10}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = P_H[7,18,24,29] \oplus C_H[7,18,24,29],$$
$$\alpha_{01}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = P_L[15] \oplus C_L[15], \text{ and}$$
$$\gamma^T \mathbf{k} = k = K_1[22] \oplus K_3[22].$$

By considering the structure of DES, it can be confirmed that these approximations are unbiased. Then considering the data mask $\alpha_{11} = \alpha_{10} + \alpha_{01}$ we obtain

$$\alpha_{11}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = P_H[7,18,24,29] \oplus P_L[15] \oplus C_H[7,18,24,29] \oplus C_L[15].$$

This data mask coincides with Matsui's best three-round linear approximation $\alpha_{11}^T \begin{pmatrix} \mathbf{p} \\ \mathbf{c} \end{pmatrix} = k$ and so it has an has imbalance 0.391 [12]. As a result, the above pair of unbiased linear approximations are entangled linear approximations with entanglement $\epsilon = 0.391$.

The experimental success rates for key bit recovery using the above pair of three-round entangled linear approximations are given in Table 2. This table also gives experimental success rates for key bit recovery for the following pair of four-round entangled linear approximations with entanglement $\epsilon = -0.122$ [12]:

$$P_H[7,18,24,29] \oplus C_L[7,18,24,29] = K_1[22] \oplus K_3[22] \oplus K_4[42,43,45,46],$$
$$P_L[15] \oplus C_H[15] \oplus C_L[27,28,30,31] = K_1[22] \oplus K_3[22] \oplus K_4[42,43,45,46].$$

The structure of DES can be used to confirm that these two approximations are unbiased. The experimental success rates in Table 2 are based on 10,000 trials.

Table 2. Success rates for key bit recovery for three- and four-round DES using entangled approximations. Results are based on 10,000 trials.

Number of plaintext-ciphertext pairs (N)	2^5	2^6	2^7	2^8	2^9	Theoretical prediction
Three-round experiment	0.6153	0.6305	0.6294	0.6301	0.6265	0.6279
Four-round experiment	0.5345	0.5432	0.5440	0.5425	0.5358	0.5388

4 Entangled Approximations and Large Data Sets

We saw in Section 3.1 that we can make an estimate for the key bit k using entangled approximations based on an asymptotic normal approximation, provided that the number N of plaintext-ciphertext pairs exceeds some bound. Suppose now that we have n such estimates for the key bit k, each based on a *packet* of N plaintext-ciphertext pairs. This gives a process for recovering the key from $T = nN$ plaintext-ciphertext pairs; given the results for n packets of size N we should clearly choose the value of the key bit k that is given by the majority of the packets.

We now determine the success rate of such a packet-based process. The probability that an individual packet gives the correct value of the key bit k is $\frac{1}{2} + \frac{\epsilon}{\pi}$. Thus the number of correct key values in a set of n packets has the following distribution

$$\mathrm{Bin}\left(n, \tfrac{1}{2} + \tfrac{\epsilon}{\pi}\right) \approx \mathrm{N}\left(\tfrac{n}{2} + \tfrac{n\epsilon}{\pi}\,;\, \tfrac{n}{4}\right).$$

Thus the probability that at least half of the n packets give the correct value for the key bit k is $1-\phi\left(-\frac{2\epsilon\sqrt{n}}{\pi}\right) = \phi\left(\frac{2\epsilon\sqrt{n}}{\pi}\right)$, where ϕ is the cumulative distribution function of a standard normal random variable. Thus over n packets of size N, where N exceeds the lower threshold established in Section 3.1, the success rate is given by $p = \phi\left(\frac{2\epsilon\sqrt{n}}{\pi}\right)$. Alternatively, for a given success rate p, we require $n = \left(\frac{\pi}{2\epsilon}\right)^2 \phi^{-1}(p)$ packets.

4.1 Experimental Confirmation

We illustrate the use of packets of texts with entangled approximations by considering the same entangled approximations for reduced-round versions of DES as we considered in Section 3.2. Table 3 gives the success rates for key bit recovery using the entangled approximations of Section 3.2 when using packets of text. Each packet contains $N = 64$ plaintext-ciphertext pairs, and the success rates are based on 10,000 trials.

Note that if we were to use packets of size one, *i.e.* $N = 1$, then our method would coincide with regular linear cryptanalysis with imbalance ϵ. However the situations of interest to us in this paper, for instance in Section 5, exclude us

Table 3. Success rates for key bit recovery for three- and four-round DES using entangled approximations and packets consisting of 64 plaintext-ciphertext pairs. Results are based on 10,000 trials.

Number of packets (n)	2^3	2^4	2^5	2^6	2^7	2^8	2^9
Three-round experiment	0.7597	0.8527	0.9270	0.9782	0.9975	1.0000	1.0000
Theoretical prediction	0.7585	0.8394	0.9197	0.9764	0.9975	1.0000	1.0000
Four-round experiment	0.5860	0.6223	0.6588	0.7231	0.8123	0.8890	0.9606
Theoretical prediction	0.5866	0.6216	0.6693	0.7321	0.8094	0.8922	0.9601

from this case. For packet sizes $2 \leq N \leq 63$ we don't have a fully-satisfactory theoretical model, but experimental results given in Table 4 confirm the following intuition. For a fixed number of texts, larger packets necessarily implies less packets and we are exploiting the underlying information less efficiently. This translates into reading down a column in Table 4 which illustrates a reduced success rate. For a given packet size, *i.e.* reading across a row in Table 4, more data translates into more packets and hence an increased success rate.

Table 4. Success rates using the entangled approximations over four-round DES when the total number of available texts is split into packets of size N. Results are based on 10,000 trials.

		Total number of texts T						
		2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}
	2^0	0.8346	0.9155	0.9741	0.9970	1.0000	1.0000	1.0000
	2^1	0.6319	0.6869	0.7512	0.8343	0.9149	0.9786	0.9966
$\uparrow$	2^2	0.6135	0.6456	0.7073	0.7769	0.8656	0.9432	0.9866
N	2^3	0.5840	0.6154	0.6465	0.7184	0.7909	0.8806	0.9490
$\downarrow$	2^4	0.5488	0.5709	0.6173	0.6694	0.7307	0.8030	0.8871
	2^5	0.5459	0.5577	0.5813	0.6215	0.6695	0.7333	0.8116
	2^6	0.5432	0.5405	0.5553	0.5860	0.6223	0.6588	0.7231

5 Separating Plaintext and Ciphertext

We can imagine situations where the connection between a plaintext and a ciphertext block is not known to the cryptanalyst. In a communication or transmission scenario, blocks might be delivered out of sequence or their ordering might be hidden or affected by some application. In another situation, cryptanalysis of an encrypted database can be practically thwarted when the correspondence between a database query (the plaintext) and its true location in the encrypted column or row (the ciphertext) is unknown. In such situations, existing linear cryptanalysis cannot be applied since analysis requires that a plaintext and its corresponding ciphertext be matched together and treated at the same time.[3] However we can still use the ideas in this paper to recover the key.

To demonstrate this we consider Matsui's best five-round approximation of DES [12]. Define the key bit

$$k = K_1[42, 43, 45, 46] \oplus K_2[22] \oplus K_4[22] \oplus K_5[42, 43, 45, 46],$$

and then Matsui's approximation, for which $\epsilon = 0.038$, is given by

$$P_H[15] \oplus C_H[15] \oplus$$
$$P_L[7, 18, 24, 27, 28, 29, 30, 31] \oplus C_L[7, 18, 24, 27, 28, 29, 30, 31] = k.$$

[3] While naïvely one might suggest that all the different plaintexts and ciphertexts be tried in all combinations, it is clear that this would be completely impractical.

Now consider launching an attack when the association between plaintext and ciphertext is lost; the attacker does not know which ciphertext corresponds to which plaintext. Perhaps we can envisage a less drastic situation where the correlation between plaintext and ciphertext is not completely lost and the attacker knows that groups of ciphertext correspond to groups of plaintexts, but inside each group the attacker is unsure which plaintext should be associated with which ciphertext. In either case regular linear cryptanalysis cannot be used since the attacker cannot compute the necessary counters. But using entangled approximations we can recover key information in both situations.

To see this, consider the two linear approximations

$$P_H[15] \oplus P_L[7, 18, 24, 27, 28, 29, 30, 31] = k, \text{ and}$$
$$C_H[15] \oplus C_L[7, 18, 24, 27, 28, 29, 30, 31] = k.$$

The first approximation does not depend on the ciphertext and the second does not depend on the plaintext. Both linear approximations are unbiased, however the two approximations are entangled since they are derived from Matsui's best five-round DES approximation.

Therefore using the techniques in this paper we can recover the key. We can do this when complete correlation is lost, by considering one packet of text as in Section 3.2, or we can use the techniques in Section 4 in the more practical situation where some packet-level correlation between plaintext and ciphertext remains. Table 5 gives the experimental success rates in the first instance, when using a varying number of plaintext-ciphertext pairs, and the success rate for this method is given by $\frac{1}{2}+\frac{\epsilon}{\pi}$. Meanwhile Table 6 confirms the predicted success rates when using the techniques of Section 4 for multiple packets, with each packet consisting of 64 plaintext-ciphertext pairs. The experimental success rates for both tables are based on 10,000 trials.

Table 5. Success rates for key bit recovery for five-round DES using entangled approximations with the separation of plaintext and ciphertext. Results are based on 10,000 trials.

Number of plaintext-ciphertext pairs (N)	2^5	2^6	2^7	2^8	2^9	Theoretical prediction
Five-round experiment	0.5092	0.5024	0.5252	0.5165	0.5127	0.5121

Table 6. Success rates for key bit recovery for five-round DES using entangled approximations with the separation of plaintext and ciphertext and packets consisting of 64 plaintext-ciphertext pairs. Results are based on 10,000 trials.

Number of packets (n)	2^3	2^4	2^5	2^6	2^7	2^8	2^9
Five-round experiment	0.5282	0.5442	0.5501	0.5762	0.6116	0.6464	0.7013
Theoretical prediction	0.5274	0.5388	0.5547	0.5771	0.6084	0.6514	0.7090

5.1 Discussion

It is important to emphasize exactly what the separation of plaintext and ciphertext entails. In both experiments, with a single packet or multiple packets, the data processing involves a plaintext counter and a ciphertext counter. The two counters are entirely separate. Thus we have an analysis using multiple linear approximations in which the plaintext and ciphertext are processed independently.

There is, of course, a small cost which can be manifested as either a modest reduction to the success rate or a moderate increase in the amount of text required for an unchanged success rate (when compared to regular linear cryptanalysis). To see the increase in text required, we note that regular cryptanalysis with a single approximation with imbalance ϵ needs

$$\left(\frac{\phi^{-1}(p)}{\epsilon}\right)^2$$

plaintext/ciphertext pairs to achieve a success rate p. Using entangled linear approximations, with n packets of sufficient large size, the number of packets needed to achieve a success rate p is

$$n = \left(\frac{\pi\phi^{-1}(p)}{2\epsilon}\right)^2.$$

As was pointed out previously, once we have a sufficiently large packet, the success rate no longer depends on the size of the packet so packets of size $N = 64$ will be sufficient. We therefore expect to attain the same success rate using separated plaintext-ciphertext instead of regular linear cryptanalysis when we have $\frac{\pi^2}{4} \times 2^6 = 2^{7.3}$ times the data. This result is confirmed experimentally for four-round DES in Table 7.

Table 7. Success rates with entangled approximations and theoretical success rates with regular linear cryptanalysis over four-round DES

Number of plaintext-ciphertext pairs (T)	$2^{1.7}$	$2^{2.7}$	$2^{3.7}$	$2^{4.7}$
Success rate (regular)	58.67	62.17	66.94	73.23

Number of plaintext-ciphertext pairs (T)	2^{9}	2^{10}	2^{11}	2^{12}
Success rate (entangled)	58.60	62.23	65.88	72.31

Note that we are not proposing that entangled approximations be used to replace regular linear cryptanalysis. Instead we are showing that there are situations where regular linear cryptanalysis cannot be used but entangled linear approximations can still give us information about the key.

6 Conclusions

In this paper we have shown that we can recover key information when using only unbiased linear approximations. In Section 5 we demonstrated a practical situation where by using multiple linear approximations in an entirely novel way—two separate approximations involving only plaintext and ciphertext bits respectively—we can extract key information when all existing techniques would fail. However, while much of the underlying analysis has been demonstrated in this paper the practical implications of entanglement still remain to be quantified. One particularly interesting situation would be using these techniques on the constituent components of ciphers. Some other directions for futher work might also include analysis in the direction of Matsui's Algorithm 2 [12], the further development of the statistical models that depend on entanglement, and an extension to linear approximations that depend on more than one key bit.

Acknowledgement

We would particularly like to thank Sean Murphy who launched this work and contributed throughout.

References

1. Biryukov, A., De Cannière, C., Quisquater, M.: On Multiple Linear Approximations. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 1–22. Springer, Heidelberg (2004)
2. Cho, J.Y.: Linear Cryptanalysis of Reduced-Round PRESENT. In: Pieprzyk, J. (ed.) CT-RSA 2010. LNCS, vol. 5985, pp. 302–317. Springer, Heidelberg (2010)
3. Collard, B., Standaert, F.-X., Quisquater, J.-J.: Improving the Time Complexity of Matsui's Linear Cryptanalysis. In: Nam, K.-H., Rhee, G. (eds.) ICISC 2007. LNCS, vol. 4817, pp. 77–88. Springer, Heidelberg (2007)
4. Collard, B., Standaert, F.-X., Quisquater, J.-J.: Experiments on the Multiple Linear Cryptanalysis of Reduced-Round Serpent. In: Nyberg, K. (ed.) FSE 2008. LNCS, vol. 5086, pp. 382–397. Springer, Heidelberg (2008)
5. Daemen, J.: Cipher and Hash Function Design. Ph.D. Thesis (March 1995)
6. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional Linear Cryptanalysis of Reduced Round Serpent. In: Mu, Y., Susilo, W., Seberry, J. (eds.) ACISP 2008. LNCS, vol. 5107, pp. 203–215. Springer, Heidelberg (2008)
7. Hermelin, M., Cho, J.Y., Nyberg, K.: A New Technique for Multidimensional Linear Cryptanalysis with Applications on Reduced Round Serpent. In: Yung, M., Liu, P., Lin, D. (eds.) ICISC 2008. LNCS, vol. 5461, pp. 383–398. Springer, Heidelberg (2009)
8. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional Extension of Matsuis Algorithm 2. In: Goos, G., Hartmanis, J., van Leeuwen, J. (eds.) Fast Software Encryption. LNCS, vol. 5665, pp. 209–227. Springer, Heidelberg (2009)

9. Junod, P.: On the complexity of Matsui's attack. In: Vaudenay, S., Youssef, A.M. (eds.) SAC 2001. LNCS, vol. 2259, pp. 199–211. Springer, Heidelberg (2001)
10. Kaliski, B.S., Robshaw, M.J.B.: Linear Cryptanalysis Using Multiple Approximations. In: Desmedt, Y.G. (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 26–39. Springer, Heidelberg (1994)
11. Kaliski, B.S., Robshaw, M.J.B.: Linear Cryptanalysis and FEAL. In: Preneel, B. (ed.) FSE 1994. LNCS, vol. 1008, pp. 249–264. Springer, Heidelberg (1995)
12. Matsui, M.: Linear Cryptanalysis Method for DES Cipher. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397. Springer, Heidelberg (1994)
13. Matsui, M.: The first experimental cryptanalysis of the Data Encryption Standard. In: Desmedt, Y.G. (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 1–11. Springer, Heidelberg (1994)
14. Murphy, S.: The Independence of Linear Approximations in Symmetric Cryptanalysis. IEEE Transactions on Information Theory 52, 5510–5518 (2006)
15. National Institute of Standards and Technology. FIPS 46-3: Data Encryption Standard (November 1998), `http://csrc.nist.gov`
16. Selçuk, A.: On Probability of Success in Linear and Differential Cryptanalysis. Journal of Cryptology 21(1), 131–147 (2008)
17. Silvey, S.D.: Statistical Inference. Chapman and Hall, Boca Raton (1975)
18. Tardy-Corfdir, A., Gilbert, H.: A Known Plaintext Attack on FEAL-4 and FEAL-6. In: Feigenbaum, J. (ed.) CRYPTO 1991. LNCS, vol. 576, pp. 172–182. Springer, Heidelberg (1992)
19. Vaudenay, S.: An Experiment on DES Statistical Cryptanalysis. In: Proceedings of the Third ACM Conference on Computer Security, pp. 386–397 (1996)