New Combined Attacks on Block Ciphers

Eli Biham[1], Orr Dunkelman[1,⋆], and Nathan Keller[2]

[1] Computer Science Department, Technion,
Haifa 32000, Israel
{biham, orrd}@cs.technion.ac.il

[2] Einstein Institute of Mathematics, Hebrew University,
Jerusalem 91904, Israel
nkeller@math.huji.ac.il

Abstract. Differential cryptanalysis and linear cryptanalysis are the most widely used techniques for block ciphers cryptanalysis. Several attacks combine these cryptanalytic techniques to obtain new attacks, e.g., differential-linear attacks, miss-in-the-middle attacks, and boomerang attacks.

In this paper we present several new combinations: we combine differentials with bilinear approximations, higher-order differentials with linear approximations, and the boomerang attack with linear, with differential-linear, with bilinear, and with differential-bilinear attacks. We analyze these combinations and present examples of their usefulness. For example, we present a 6-round differential-bilinear approximation of s^{5}DES with a bias of 1/8, and use it to attack 8-round s^{5}DES using only 384 chosen plaintexts. We also enlarge a weak key class of IDEA by a factor of 512 using the higher-order differential-linear technique. We expect that these attacks will be useful against larger classes of ciphers.

1 Introduction

In a differential attack [5], the attacker seeks a fixed input difference that propagates through the nonlinear parts of the cipher to some fixed output difference with usually high (or zero) probability. Such pair of differences with the corresponding probability is called a *differential*. In the attack, the attacker asks for the encryption of pairs of plaintexts with the input difference given by the differential, and checks whether the output difference predicted by the differential occurs (with the predicted probability).

In a linear attack [30], the attacker seeks a linear approximation between the parity of a subset of the plaintext bits and the parity of a subset of the ciphertext bits with a biased probability. The attacker asks for the encryption of many plaintexts, and checks whether the linear relation predicted by the linear approximation is satisfied or not.

[⋆] The research presented in this paper was supported by the Clore scholarship programme.

H. Gilbert and H. Handschuh (Eds.): FSE 2005, LNCS 3557, pp. 126–144, 2005.

In 1994, Langford and Hellman [28] showed that both kinds of cryptanalysis can be combined together by a technique called *differential-linear cryptanalysis*, in which the differential is used to obtain a linear approximation (between two encryptions) with bias 1/2. The technique was improved in [8, 27], allowing the usage of differentials with probability lower than 1, thus making the technique applicable to a larger set of block ciphers.

The differential-linear technique was applied to analyze several (reduced versions of) block ciphers, such as: DES [32] (attacked in [28, 8]), IDEA [26] (attacked in [13, 20]), Serpent [1] (attacked in [9]), and COCONUT98 [35] (attacked in [8]). Some of the attacks are the best known attacks against the respective versions of the ciphers. It was also shown that the ciphertext-only extensions of differential and linear cryptanalysis work with differential-linear cryptanalysis as well [10].

Langford and Hellman's technique is an example for devising the distinguisher (to be used in the attack) as a combination of two much simpler parts. In this case, a combination of a differential and a linear approximation. Such combinations were later used in other cryptanalytic techniques, e.g., cryptanalysis using impossible differentials [6, 7] (miss in the middle), and boomerang attacks [36], both using combinations of differentials.

In this paper we present several new combinations of the differential, the higher-order differential, the boomerang, the linear, and the bilinear techniques. All of these combinations treat the distinguished part of the cipher as a cascade of two (or even three) sub-ciphers.

First, we show how to combine the differential cryptanalysis with the bilinear cryptanalysis [14]. Bilinear cryptanalysis is a generalization of linear cryptanalysis specially designed for Feistel block ciphers. In bilinear cryptanalysis the attacker studies relations between bilinear functions of the bits of the plaintext and bilinear functions of the bits of the ciphertext. Usually, the results of bilinear cryptanalysis are comparable with those of ordinary linear cryptanalysis. However, there are ciphers that are relatively strong against linear cryptanalysis but are vulnerable to bilinear cryptanalysis. For example, s^{5}DES [21] is stronger than DES against linear cryptanalysis while the best 3-round bilinear approximation of s^{5}DES has a bias of 1/4, which is much larger than the corresponding linear approximation for DES.

We show that bilinear approximations can be combined with differentials essentially in the same way as ordinary linear approximations are combined. However, there are some differences between a regular differential-linear attack and a differential-bilinear attack. We explore the similarities and the differences between the two attacks, and apply the differential-bilinear technique to attack 8-round s^{5}DES.

The next combination we discuss is the higher order differential-linear attack. Higher-order differential cryptanalysis [2, 22, 25] is a generalization of differential cryptanalysis that uses differentials of more than two plaintexts. In the higher-order differential attack the attacker analyses the development of the XOR of the intermediate data during the encryption of a set of plaintexts satisfying some conditions. Attacks which resemble higher-order differential attack, such as SQUARE-like attacks [12, 18, 24, 29], can also be combined with linear cryptanalysis.

We show that higher-order differentials (and SQUARE-like properties) can also be used as a building block in a two-phase attack. In higher-order differential-linear cryptanalysis, the attacker examines sets of plaintexts that have the input difference of the higher-order differential. The higher-order differential predicts the XOR value of all the intermediate encryption value after the higher-order differential. Then, the linear approximation can be applied to the entire set to predict the parity of a subset of the ciphertext bits (of all the ciphertexts).

The data complexity of the higher-order differential-linear attack is proportional to $2^{2m}/p^2q^{2m}$, where p is the probability of the higher-order differential, q is the bias of the linear approximation, and m is the number of plaintexts in each set. Therefore, the attack can be used only if either the structure is small enough or the linear approximation is very good (e.g., with bias 1/2). Such instances can occur in block ciphers, especially in weak key classes for which very strong and unexpected properties hold. For example, in the linear weak key class of IDEA [17], a specially built approximation has a bias of 1/2. We show that in the case of IDEA the size of the linear weak key class is increased from 2^{23} keys in the class of a regular linear attack to 2^{32} keys using a higher-order differential-linear attack.

The last combination we discuss in this paper is the differential-linear boomerang technique. The boomerang attack [36] treats the cipher as a cascade of two sub-ciphers, and exploits two differentials, one for each sub-cipher, in order to obtain some information on the differences using an adaptive chosen plaintext and ciphertext process. In a differential-linear boomerang attack, the attacker constructs a pair of encryptions whose difference in the intermediate encryption value is known by means of the boomerang technique. This pair can then be analyzed by means similar to those of the differential-linear cryptanalysis. Moreover, it appears that the linear boomerang is a special case of a more general attack. By decomposing the first sub-cipher into two sub-sub-ciphers (and the cipher into three sub-ciphers in total), we can apply the differential-linear (or the differential-bilinear) attack to the cipher.

One interesting feature of the (differential-)(bi)linear boomerang attack is that this is the first attack that treats the cipher as a cascade of three sub-ciphers successfully, while all previous works treat the cipher as a cascade of at most two sub-ciphers.

The paper is organized as follows: In Section 2 we shortly sketch the basic differential-linear attack. In Section 3 we present differential-bilinear cryptanalysis and apply it to DES and s^{5}DES. In Section 4 we discuss higher-order differential-linear cryptanalysis and present several applications of the attack, including increasing the linear weak class of IDEA. In Section 5 we introduce (differential-)(bi)linear boomerang attacks. This set of attacks are combinations of the boomerang technique with the (differential-)(bi)linear attack. We concentrate on the differential-bilinear boomerang attack, as this attack is the most general one (while the other variants can be treated as special cases of this attack). Finally, Section 6 concludes this paper.

2 Preliminaries

2.1 Notations

We use notations based on [3, 5] for differential and linear cryptanalysis, respectively. In our notations Ω_P, Ω_T are the input and the output differences of the differential, and λ_P, λ_C are the input and the output subsets (denoted by bit masks) of the linear approximation. We also use λ_T to denote the input subset in some cases.

Let $E = E_1 \circ E_0$ be a block cipher, i.e., $C = E_k(P) = E_{1_k}(E_{0_k}(P))$. For example, if E is DES, then E_0 can be the first eight rounds of DES, while E_1 are the last eight rounds. For sake of simplicity, we omit the key, as it is clear that encryption is done using a secret key. We denote the partial encryption of P (and the partial decryption of C) by T, i.e., $T = E_0(P) = E_1^{-1}(C)$.

The last notation is the scalar product of two strings x and y and is denoted by $x \cdot y$.

2.2 Differential-Linear Cryptanalysis

Langford and Hellman [28] show that a concatenation of a differential and a linear approximation is feasible. The main idea in the combination is to encrypt pairs of plaintexts, and check whether the corresponding ciphertext pairs have the same parity of the output mask or not.

Let $\Omega_P \rightarrow \Omega_T$ be a differential of E_0 with probability 1. Let $\lambda_T \rightarrow \lambda_C$ be a linear approximation of E_1 with bias $\pm q$. We start with a pair of plaintexts P_1 and $P_2 = P_1 \oplus \Omega_P$. After the partial encryption through E_0, the intermediate encryption values are T_1 and $T_2 = T_1 \oplus \Omega_T$, respectively. For any intermediate encryption value T and its corresponding ciphertext C, $\lambda_T \cdot T = \lambda_C \cdot C$ with probability $1/2 + q$. Therefore, each of the relations $\lambda_C \cdot C_1 = \lambda_T \cdot T_1$ and $\lambda_C \cdot C_2 = \lambda_T \cdot T_2 = \lambda_T \cdot T_1 \oplus \lambda_T \cdot \Omega_T$ is satisfied with probability $1/2 \pm q$. Hence, with probability $1/2 + 2q^2$ the relation $\lambda_C \cdot C_1 = \lambda_C \cdot C_2 \oplus \lambda_T \cdot \Omega_T$ holds.

We note that λ_T and Ω_T are known, and thus, we have constructed a condition on C_1 and C_2 which has probability $1/2 + 2q^2$, while for a random pair of ciphertexts, this condition is satisfied with probability $1/2$. This fact can be used in distinguishers and in key recovery attacks. Hellman and Langford also noted that it is possible to use truncated differentials [22] as long as $\lambda_T \cdot \Omega_T$ is predictable.

As both difference and parity are linear operations, the two linear approximations in E_1 in both encryptions can be combined into an approximation of E of the form

$$E_{1_1}\text{–differential–}E_{1_2},$$

where the lower subscript denotes whether the sub-cipher is in the first encryption or in the second, and "differential" refers to the differential combiner that ensures that the parities of the data before transition from E_0 to E_1 in both encryptions are always equal (or always differ).

This led to the introduction of a differential-linear approximation for 6-round DES which was composed of a 3-round differential and a 3-round linear approximation. The differential-linear approximation was then used to attack 8-round DES. The attack requires 768 chosen plaintexts, and has the lowest data requirements between all attacks on 8-round DES.

Later research [8, 27] showed that it is possible to have $\lambda_T \cdot \Omega_T$ unknown but fixed. Also, it was shown that when the differential-linear technique is applicable when the differential has probability $p \neq 1$. In that case the probability that $\lambda_T \cdot T_1 = \lambda_T \cdot T_2 \oplus \lambda_T \cdot \Omega_T$ is $1/2 + p'$, where $p' = p/2$, and thus the event $\lambda_C \cdot C_1 = \lambda_C \cdot C_2 \oplus \lambda_T \cdot \Omega_T$ holds with probability $1/2 + 4p'q^2 = 1/2 + 2pq^2$.

As we demonstrate later, in some of the attacks that we present this property does not hold. That is, the attacker has to know the exact value of the difference Ω_T, and in some cases, only certain values of the difference Ω_T can be used in the combined attack.

Moreover, even if $\Omega_T \cdot \lambda_P$ is unknown to the attacker but constant for a given key, the attack still succeeds. In that case we know that the value $\lambda_C \cdot C_1 \oplus \lambda_C \cdot C_2$ is either 0 or 1, with a bias of $2q^2$. This case is similar to the case in linear cryptanalysis, when $\lambda_K \cdot K$ is unknown, and can be either 0 or 1.

3 Differential-Bilinear Attack

3.1 Bilinear Cryptanalysis

The bilinear attack [14] is a generalization of linear cryptanalysis aimed at Feistel ciphers. The attack considers approximations involving bilinear terms of the input, the output, and the key. The reason this attack aims at Feistel ciphers is that it is easier to find such bilinear approximations for Feistel ciphers.

For the description of the bilinear approximations we adopt the notations used in [14]. We also put aside the probabilistic nature of some of the steps for sake of clarity (of course, when we use the approximations we take the probabilities back into account). Let the input value of the r-th round in a Feistel cipher be $(L_r[0, 1, ..., n-1], R_r[0, 1, ..., n-1])$, where L stands for the left half of the data and R stands for the right half (note that R_0 and L_0 compose the plaintext). Furthermore, we denote the input and the output values of the F-function in the r-th round by $I_r[0, 1, ..., n-1]$, and $O_r[0, 1, ..., n-1]$, respectively. Due to the structure of a Feistel cipher $I_r = R_r$, $R_{r+1} = L_r \oplus O_r$, and $L_{r+1} = R_r$.

Let α be a subset of $\{0, 1, ..., n-1\}$, then $L_r[\alpha] = \oplus\{L_r[s] | s \in \alpha\} = \oplus_{s\in\alpha} L_r[s]$, i.e., $L_r[\alpha]$ is the parity of all bits in the left half masked by α. Similarly $R_r[\beta]$ is the parity all bits in the right half masked by β.

According to the Feistel round, for any mask α, β and any round r:

$$L_{r+1}[\beta] \cdot R_{r+1}[\alpha] \oplus R_r[\beta] \cdot L_r[\alpha] = I_r[\beta] \cdot O_r[\alpha].$$

Such 1-round bilinear approximations can be concatenated to obtain bilinear approximations of several rounds. Concatenation requires some additional conditions, and also introduces some probability to the whole approximation. We note that in some cases the relations involve key bits in bilinear terms as well, e.g., $L_r[12] \cdot K_r[15]$. One-round approximations can also be extended such that they include linear terms in addition to the bilinear ones. In this case, the concatenation is more complex and can be achieved only if the linear terms fulfill some additional requirements. The full description of bilinear approximations is given in [14]. The general form of the obtained bilinear approximation is

$$\begin{gathered} L_0[\alpha_0] \cdot R_0[\beta_0] \oplus R_0[\gamma_0] \oplus L_0[\delta_0] \oplus L_n[\alpha_n] \cdot R_n[\beta_n] \oplus R_n[\gamma_n] \oplus L_n[\delta_n] = \\ L_0[\epsilon_0] \cdot K[\epsilon_1] \oplus R_0[\zeta_0] \cdot K[\zeta_1] \oplus L_n[\eta_0] \cdot K[\eta_1] \oplus R_n[\theta_0] \cdot K[\theta_1] \oplus K[\iota_1] \end{gathered} \tag{1}$$

where K is the key (or more precisely, the list of subkeys), and all Greek letters represent some mask.

Given the above approximation, the bilinear attack resembles the linear attack. Many plaintext/ciphertext pairs are gathered, and for any guess of $K[\epsilon_1]$, $K[\zeta_1]$, $K[\eta_1]$, $K[\theta_1]$, and $K[\iota_1]$, the attacker counts how many pairs satisfy the approximation. The guess for which the above approximation holds with the expected probability of $1/2 + q$ is assumed to be the right guess.

We note that in a bilinear approximation there might be bilinear expressions involving the subkey. This fact has implications on the differential-bilinear attack which we explore later.

3.2 Differential-Bilinear Cryptanalysis

Roughly speaking, the differential-bilinear attack encrypts many pairs of plaintexts, and examines Whether the obtained pair of ciphertexts satisfy some bilinear approximation or not. This is very similar to the way that differential and linear cryptanalysis are combined.

We shall assume, without loss of generality, that the bilinear approximation has the form presented in Equation (1), and that the probability of the approximation is $1/2 + q$. We note that it is possible to have several bilinear terms in the approximation, but this fact does not change our analysis. We denote the differential to be concatenated by $\Omega_P \rightarrow \Omega_T$, and assume that the differential has probability p.

The attacker chooses pairs of plaintexts P_1 and $P_2 = P_1 \oplus \Omega_P$. With probability p the the intermediate encryption values T_1 and T_2, respectively, have a difference that satisfies the equality

$$T_{1L}[\alpha_0] \cdot T_{1R}[\beta_0] \oplus T_{1L}[\gamma_0] \oplus T_{1R}[\delta_0] = T_{2L}[\alpha_0] \cdot T_{2R}[\beta_0] \oplus T_{1L}[\gamma_0] \oplus T_{2R}[\delta_0], \tag{2}$$

where T_{iL} is the left half of T_i, and similarly T_{iR} is the right half of T_i. We note that under the random distribution[1] assumption, in the $(1-p)$ of the cases where

[1] We note that whether this assumption holds for a given cipher needs to be throughly investigated, and if possible verified as done in [9].

the differential does not hold, Equation (2) holds in half of the times. Thus, the probability that Equation (2) holds is $p+(1-p)/2 = 1/2+p/2$, and the bias is $p' = p/2$.

Then, similarly to the differential-linear case, the pair of ciphertexts C_1 and C_2 satisfies the following equation

$$C_{1L}[\alpha_n]\cdot C_{1L}[\beta_n]\oplus C_{1L}[\gamma_n]\oplus C_{1R}[\delta_n] = C_{2L}[\alpha_n]\cdot C_{2R}[\beta_n]\oplus C_{2L}[\gamma_n]\oplus C_{2R}[\delta_n] \quad (3)$$

with probability $1/2+4p'q^2 = 1/2+2pq^2$.

However, unlike differential-linear cryptanalysis where any differential can be used for the combined attack, in the bilinear case the situation is more complicated. This is due to the fact that bilinear approximations require more knowledge about the data than linear approximations. In some cases, the required information is not given by the differential.

It appears that the knowledge of the difference $L_{T_1}[\alpha_0]\oplus L_{T_2}[\alpha_0]$ and $R_{T_1}[\beta_0]\oplus R_{T_2}[\beta_0]$ in the two encryptions does not imply the knowledge of the difference between the $L_T[\alpha_0]\cdot R_T[\beta_0]$ values. Thus, the attacker is restricted to the cases where the knowledge suggested by the difference Ω_T suffices to know the difference of the $L_T[\alpha_0]\cdot R_T[\beta_0]$ values. This is clearly the case when $\alpha\cdot\Omega_{TL} = \beta\cdot\Omega_{TR} = 0$, i.e., if the parity of the differences in the bits masked by α and β is zero. Another example is when there are six active bits in the output of the differential a, b, c, d, e and f, and the bilinear approximation is $a\cdot b+c\cdot d+e\cdot f+a\cdot f+c\cdot b+e\cdot d$. For an arbitrary bilinear relation $\sum_{\alpha,\beta} L_{T_i}[\alpha]\cdot R_{T_i}[\beta]$, where α and β are masks, the difference between the two sums can be predicted (to be zero) whenever the following two conditions hold simultaneously: (1) Each $L_{T_i}[\alpha]$ appears an even number of times in products with $R_{T_i}[\beta]$'s whose difference is 1, and (2) Each $R_{T_i}[\beta]$ appears an even number of times in products with $L_{T_i}[\alpha]$'s whose difference is 1.

We note that the linear terms of the approximation behave in the same way as in differential-linear cryptanalysis. This is due to the way the attack works — the attacker examines the difference in the output mask of two encryptions, and as long as the linear terms do not affect the bias of the difference in the output mask, the linear terms do not change the attack.

A more formal way to describe a differential-bilinear approximation is: Assume that the cipher E can be decomposed to two sub-ciphers $E = E_1\circ E_0$, where the differential $\Omega_P \to \Omega_T$ (and probability p) is used in E_0, and a bilinear approximation is used for E_1. Also assume that the bits predicted in Ω_T are sufficient to know the difference in the $L_T[\alpha_0]\cdot R_T[\beta_0]$ values with bias $p/2$. Let b_1 and b_2 denote the outputs of the bilinear approximation in the first and the second encryptions, respectively. The combination between the differential and the bilinear approximation can be represented by the following extended bilinear approximation:

$$b_1\text{–differential–}b_2,$$

where "differential" refers to the differential combiner. A distinguishing attack or a key recovery attack based on the differential-bilinear property is similar to an ordinary differential-linear attack — the attacker encrypts many plaintext pairs, and checks in how many of the pairs satisfy Equation (3).

The probability that a pair of ciphertexts (C_1, C_2), originating from a pair of plaintexts $(P_1, P_2 = P_1 \oplus \Omega_T)$, to satisfy Equation (3) is $1/2+4p'q^2 = 1/2+2pq^2$.

An interesting fact that will be demonstrated in the bilinear approximation of DES is that the subkey may be a part of the bilinear approximation. While in a linear approximation the linear factors of the key are independent of the plaintext (or the ciphertext), and can be treated like such, in a bilinear approximation the key may have a bilinear term involving the plaintext (or the ciphertext). Thus, Equation (3) might involve unknown key terms. When the equation involves unknown key terms, the attacker has to try all possible combinations for these key terms in the attack.

3.3 Applying Differential-Bilinear Cryptanalysis to DES and to s^{5}DES

In [14] a 3-round bilinear approximation of DES is presented. The approximation has a bias of $q = 1.66 \cdot 2^{-3}$ which is slightly better than the best 3-round linear approximation (that has a bias of $1.56 \cdot 2^{-3}$). The bilinear approximation is as follows:

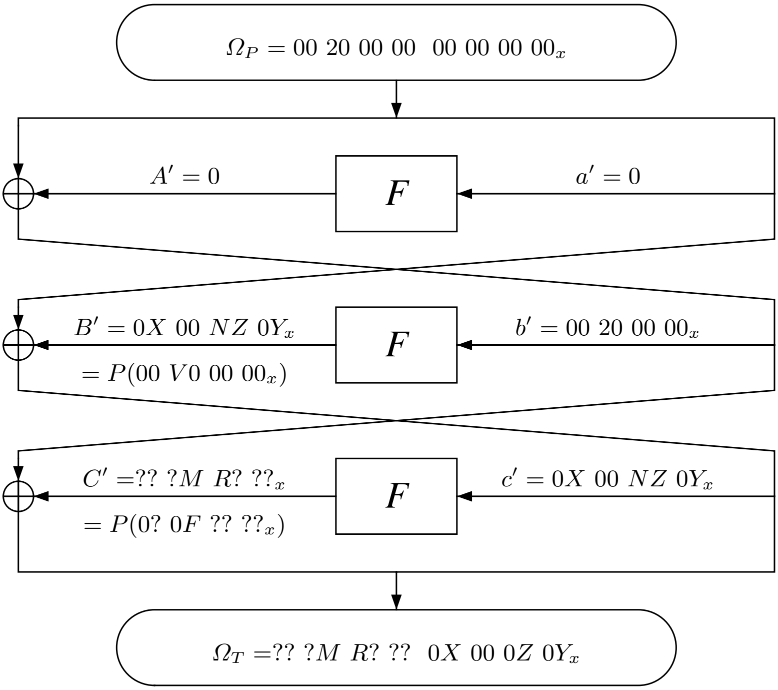
$$\begin{gathered} L_0[3,8,14,25] \oplus R_0[17] \oplus L_0[3] \cdot R_0[16,17,20] \oplus \\ L_3[3,8,14,25] \oplus R_3[17] \oplus L_3[3] \cdot R_3[16,17,20] \\ = K[sth] \oplus L_0[3] \cdot K[sth'] \oplus L_3[3] \cdot K[sth''], \end{gathered}$$

where (L_0, R_0) is the plaintext (or in our case the intermediate encryption value), (L_3, R_3) is the ciphertext, and $K[sth], K[sth']$, and $K[sth'']$ are subsets of the key bits.

We can concatenate the above bilinear approximation to a differential that predicts a zero difference in $L_0[3] \cdot R_0[16,17,20]$. The best 3-round differential that satisfies the requirements for concatenating the differential and the bilinear parts is presented in Figure 1. It has probability 46/64, and has the following structure: The first round has a zero input difference. The second round has an input difference with one active S-box — $S3$. The input difference of 4_x to $S3$ may cause an output difference whose bit 2 (of $S3$) is inactive with probability 28/64. If this is the case, then the masked bits of the input of the bilinear approximation are guaranteed to have a zero difference after the third round. Otherwise (with probability 36/64), bit 2 of the output of $S3$ is active. This bit enters $S4$ in the third round, and with probability 1/2 the output difference of $S4$ does not affect the bits masked by the input mask of the bilinear approximation, and thus, with probability $28/64 + 1/2 \cdot 36/64 = 46/64$ a pair with input difference $\Omega_P = (0_x, 00\ 20\ 00\ 00_x)$ has a zero difference in Ω_T in the bits masked by the bilinear approximation.

According to the previous analysis, the bias of the 6-round differential-bilinear approximation that starts with the above input difference is

$$2pq^2 = 2\frac{46}{64}(1.662^{-3})^2 = 1.98 \cdot 2^{-5}.$$



(where $X, Y \in \{0, 4\}$, $Z \in \{0, 1\}$, $M \in \{0, 2, 4, \ldots, E_x\}$, $R \in \{2, 4, 6\}$, $F \in \{0, 1, 2, 3, 8, 9, A_x, B_x\}$, $N \in \{0, 8\}$, $V \in \{3, 5, 6, 7, 9, A_x, B_x, C_x, D_x, E_x, F_x\}$ and where ? is any arbitrary value.)

Fig. 1. A 3-Round Differential of DES with Probability 46/64

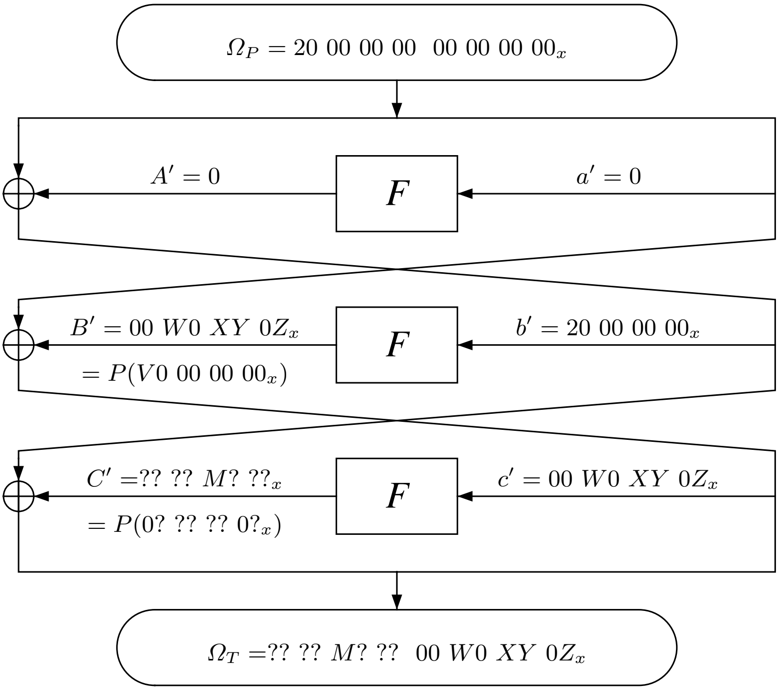
This bias is slightly lower than the bias of the best 6-round differential-linear approximation (that equals to $2.43{\cdot}2^{-5}$), and thus, the differential-bilinear attack on 8-round DES requires more data than the corresponding differential-linear attack.

An example that illustrates the advantages of the differential-bilinear cryptanalysis over a regular differential-linear attack is s^{5}DES [21]. In [15] the following bilinear approximation with bias $q = 1/4$ is presented:

$$L_0[17,23,31] \oplus R_0[1,5] \oplus L_0[9] \cdot R_0[5] \oplus$$
$$L_3[17,23,31] \oplus R_3[1,5] \oplus L_3[9] \cdot R_3[5] = K[sth],$$

where $K[sth]$ is a subset of the key bits. This bilinear approximation can be concatenated to the 3-round differential with probability 1 presented in Figure 2. The differential assures that the difference in the input bits of the bilinear term of the bilinear approximation is zero with probability 1. Thus, the bias of the differential-bilinear approximation is:

$$2pq^2 = 2(1/4)^2 = 1/8$$



(where $V \in \{1, \ldots, F_x\}$, $W \in \{0, 8\}$, $X \in \{0, 8\}$, $Y \in \{0, 2\}$, $Z \in \{0, 2\}$, $M \in \{0, \ldots, 7\}$, and ? is any arbitrary value)

Fig. 2. A 3-Round Differential of s^{5}DES with Probability 1

This differential-bilinear approximation can be used to attack 8-round s^{5}DES using 384 chosen plaintexts and time complexity of $2^{20.2}$ encryptions. The attack finds about 90 suggestions for 16 bits of the key, where the right value is among the suggested values with probability of 65.5%.

4 Combining Higher-Order Differential and Linear Attacks

4.1 Higher-Order Differential Cryptanalysis and SQUARE-Like Attacks

Higher-order differential cryptanalysis [2, 22, 25] is a generalization of differential cryptanalysis that exploits the algebraic structure of the cipher. In a higher-order differential attack the attacker asks for the encryption of a structured set of chosen plaintexts and analyses the XOR value (or some other function) of the ciphertexts. The motivation of the attack is the fact that while it is well known that linear relations between sets of bits during encryption should be avoided, in some instances higher-order relations between sets of bits can be found.

Ordinary differential cryptanalysis resembles an examination of the derivative of the nonlinear function of the cipher. It seeks cases with high enough probability in which the nonlinear function can be approximated by a linear function. Similarly, higher-order differential cryptanalysis looks at the higher-order derivatives of the nonlinear function and seeks cases where the derivatives can be predicted with high probability.

A close relative of the higher-order differential attack is the class of the SQUARE-like attacks [12, 18, 24, 29]. These attacks are aimed against ciphers in which small portions of the bits are interleaved by a strong nonlinear function while the main interleaving stage is linear. This is the case in many of the SP networks being in use today, and in particular in the AES. In this kind of attacks, the attacker examines a set of plaintexts, chosen such that the input to one of the non-linear part gets all the possible values. Thus, the attacker knows that the set contain all the intermediate values (after the nonlinear stage), but she does not know which value has originated from which plaintext. In this case, the attacker does not look for the XOR of the ciphertexts, but rather for more complicated functions, such as whether each of the possible values appears only once or not. SP networks with only a few rounds are especially vulnerable, as very efficient attacks can be devised, no matter what the non-linear function is [12].

Both higher-order differential cryptanalysis and SQUARE-like attacks, start with a set of specially chosen plaintexts, and look for some special structure in the obtained set of ciphertexts. The difference between the two attacks is the form of the special structure we expect/look for in the ciphertexts set.

4.2 The Higher-Order Differential-Linear Attack

The combination of higher-order differentials with linear approximations is similar to ordinary differential-linear cryptanalysis. The attacker uses the higher-order differential (or the SQUARE property) to predict the XOR value of the sets of masked bits in all of the elements of the structure, and then uses the linear approximation to compare this value with the XOR of the masked ciphertext bits in all of the encryptions.

Let *Set* be a set of plaintexts $\{P_1, P_2, \ldots, P_m\}$ such that the higher-order differential predicts (with some probability p) the value $\oplus_{i=1}^{m} T_i$ where the T_i's are the intermediate encryption values. Under standard independence assumptions, this means that the parity of any subset of bits taken over all intermediate encryption values is biased with a bias of $p' = p/2$. We also assume that there is a linear approximation that predicts the value of $\lambda_T \cdot T \oplus \lambda_C \cdot C$ with probability $1/2 + q$.

Lemma 1. *Let the event I be*

$$I = \{\lambda_P \cdot (T_1 \oplus \ldots \oplus T_m) = \lambda_C \cdot (C_1 \oplus \ldots \oplus C_m)\}.$$

Then (under standard independence assumptions) $\Pr[I] = 1/2 + 2^{m-1}q^m$.

Before the proof we note that I is actually the event that the XOR of the input mask, taken over all intermediate encryption values, is equal to the XOR of the output mask, taken over all ciphertexts.

Proof. The proof of the lemma is by induction on m, and is very similar to the proof of Matsui's Piling-up Lemma [30]. If $m = 1$, there is only one approximation and thus the probability equals to $1/2 + q$. Assume that the claim holds for structures of size k and consider a structure of size $k + 1$. We divide the structure into two structures, one consisting of k ciphertexts, and the other consisting of one ciphertext. The division into two structures can be done at random. Consider the probabilities of the events I in the two structures, i.e., consider each structure as an independent structure and consider the probability of the events I corresponding to these new structures. Clearly, the event I occurs for the whole structure if and only if the corresponding events I_1, I_k occur either for both structures or for none of them. By the induction hypothesis, the probability of such an event equals to:

$$(1/2 + 2^{k-1}q^k)(1/2 + q) + (1/2 - 2^{k-1}q^k)(1/2 - q) =$$
$$1/4 + 2^{k-2}q^k + 2^{k-1}q^{k+1} + q/2 + 1/4 - 2^{k-2}q^k + 2^{k-1}q^{k+1} - q/2 = 1/2 + 2^k q^{k+1}$$

Thus, by induction, the lemma is proven. Q.E.D.

Lemma 2. *Given a set of plaintexts with the input requirements of the higher-order differential, the bias of the event that the XOR of the output mask in all the ciphertexts equal to the value predicted by the linear approximation is*

$$\hat{b} = 2^{m-1}pq^m. \tag{4}$$

Proof. The proof is a combination of the result of the previous lemma with the probability of the higher-order differential. Let Z_1, Z_2 be the boolean variables defined as $Z_1 = \lambda_P \cdot (T_1 \oplus ... \oplus T_m)$, and $Z_2 = \lambda_C \cdot (C_1 \oplus ... \oplus C_m)$. We are interested in the probability $P(Z_2 = 0)$. If this probability differs from 1/2, then we can use this property for the attack. Combining the higher-order differential with the results on the linear approximation obtained above, we get that $P(Z_1 = 0) = 1/2 + p/2$ and $P(Z_1 = Z_2) = 1/2 + 2^{m-1}q^m$. Therefore,

$$P(Z_2 = 0) = P(Z_1 = 0) \cdot P(Z_2 = Z_1) + P(Z_1 = 1) \cdot P(Z_2 \neq Z_1) =$$
$$(1/2 + p/2)(1/2 + 2^{m-1}q^m) + (1/2 - p/2)(1/2 - 2^{m-1}q^m) = 1/2 + 2^{m-1}pq^m.$$

Q.E.D.

Note that differential-linear cryptanalysis can be considered as a special case of higher-order differential-linear cryptanalysis, where the size of the structure is 2. Using Formula (4), the bias of the approximation is $\hat{b} = 2pq^2$.

4.3 Applications of Higher-Order Differential-Linear Cryptanalysis

Our first application of the higher-order differential-linear cryptanalysis is a generic attack. Let E be a Feistel block cipher with a bijective round function F. Denote the block size of E by $2n$. Assume that E has an r-round linear approximation with bias 1/2. We combine this r-round linear approximation with a 3-round higher-order differential that exists with probability 1 for all such ciphers.

Let a word that is constant for all plaintexts in the structure be denoted by C. Let a word that assumes all possible values (a permutation) for a given structure be denoted by P, and let a word in which the XOR value of all the plaintexts in the structure is zero be denoted by B. For example (P, P) is a structure of 2^n plaintexts, where every possible value of the left half appears once, as well as every possible value of the right half (and we assume no relation between these instances). Another example is (B, C) — a structure of 2^n plaintexts where the right half is fixed in all the plaintexts, and the XOR of all the values in the left half is zero.

For the Feistel cipher described above, the following 3-round higher-order differential holds with probability 1:

$$(P, C) \xrightarrow{F} (C, P) \xrightarrow{F} (P, P) \xrightarrow{F} (P, B).$$

(This kind of property was first used in [4] with different attack methods). As can be seen from the higher-order differential, the attacker knows for certain that the XOR of the texts in the structure at the end of round 3 is 0, and the same is true for the XOR value in any specific bit as well. The 3-round higher-order differential can be combined with the linear approximation to devise a $(k + 3)$-round higher-order differential-linear approximation of the cipher. The overall bias of the approximation is 1/2, and thus the approximation requires several structures of 2^n chosen plaintexts to distinguish between the cipher and a random permutation.

This generic attack can be applied to FEAL [33]. FEAL is a 64-bit Feistel block cipher, with a bijective round function. There exists a linear approximation for three rounds of the cipher with bias 1/2 (see [31] for details). We can combine this linear approximation with the 3-round higher-order differential to devise a 6-round higher-order differential-linear approximation with bias 1/2 (and a set size of 2^{32} plaintexts), and use it to distinguish between FEAL-6 and a random permutation. This distinguisher can be used in a key recovery attacks on FEAL-7 and FEAL-8. Even though these attacks are far from being the best known attacks, they demonstrate the feasibility of higher-order differential-linear cryptanalysis.

Another application of this technique is a weak key class of the block cipher IDEA [26]. IDEA has a 64-bit block size and it consists of 8.5 rounds. It is based on operations on four words of 16-bit each.

There is a weak key class of 2^{32} keys, each having zero in 96 positions, that can be detected using a higher-order differential-linear attack. The underlying linear approximation is the one used in the linear weak key class of IDEA of 2^{23} keys in [17]. The approximation has bias 1/2, and it propagates through IDEA by exploiting the fact that for the weak key class the multiplication operation can be approximated with bias 1/2.

Our weak key class uses a 3-round higher-order differential that starts with sets of the form (P, C, P, C), for which after three rounds the XOR of the least significant bits of the first and the second words are zero. The linear approximation is used in the remaining 5.5 rounds, and it has a bias of 1/2. Thus, for this weak key class, the output mask of all ciphertexts in a given set is the same. We can use this fact and about 100 sets to identify whether the key used in the encryption is in the weak key class.

We conclude that our new weak key class contains 2^{32} keys, 512 times more keys than the original linear weak key class. The membership tests requires about 2^{23} chosen plaintexts with a negligible amount of computation time.

We conclude that the higher-order differential-linear attack is feasible, and that in some cases it can be used to improve existing attacks and to devise new attacks. At this stage we have not found a published cipher for which our new technique yields the best attack, even though it is clear that one can easily "engineer" a dedicated cipher with this property.

4.4 Related Work

We first note that the higher-order differential-linear attack was developed independently in [34] under the name square-nonlinear attack. The attack combines a SQUARE property with a nonlinear approximation whose input is linear. Thus, the analysis can be reproduced, and despite the non-linear nature of the attack, the biases behave in the same way. The square-nonlinear attack was used to attack reduced round version of SHACAL-2.

Another related work is the chosen plaintext linear attack [23]. In the chosen plaintext linear attack, the attacker encrypts structures of plaintexts, chosen such that the input mask is the same for all values in the structure. An alternative description would say that the set is chosen such that the difference of the intermediate encryption values is 0 in the bits considered by the approximation. In such a case the attacker can examine only the output parities. This method can be used to either eliminate rounds from the approximation, or to reduce the number of candidate subkeys (as rounds before the approximation no longer play an active role in determining whether the approximation holds or not).

While there are similarities between the chosen plaintext linear attack and our higher-order differential-linear attack, there are also major differences. Our proposed technique looks for the XOR of all ciphertexts in the set, while the chosen plaintext linear attack examines the approximation in each ciphertext separately.

Actually, chosen plaintext linear attack will usually lead to a better attack, as it takes into consideration each plaintext/ciphertext pair, rather than performs an operation that "cancels" the information conveyed in 2^{16} (or even more) plaintext/ciphertext pairs. On the other hand, the chosen plaintext linear attack fixes bits of the plaintext, leading to a smaller number of possible plaintext/ciphertext values. Another advantage of our attack is its ability to "correct" wrong structures, i.e., assume that the input mask is biased with some probability (rather than fixed).

5 Combining the Boomerang Attack with Linear and Bilinear Techniques

5.1 The Boomerang Attack

The main idea behind the boomerang attack [36] is to use two short differentials with relatively high probabilities instead of one long differential with very low probability. The attack treats the block cipher $E : \{0,1\}^n \times \{0,1\}^k \to \{0,1\}^n$ as a cascade $E = E_1 \circ E_0$, such that for E_0 there exists a differential $\alpha \to \beta$ with probability p_0, and for E_1 there exists a differential $\gamma \to \delta$ with probability p_1. The distinguisher performs the following boomerang process:

- Ask for the encryption of a pair of plaintexts (P_1, P_2), such that $P_1 \oplus P_2 = \alpha$, and denote the corresponding ciphertexts by (C_1, C_2).
- Calculate $C_3 = C_1 \oplus \delta$ and $C_4 = C_2 \oplus \delta$, and ask for the decryption of the pair (C_3, C_4). Denote the corresponding plaintexts by (P_3, P_4).
- Check whether $P_3 \oplus P_4 = \alpha$.

We denote the intermediate encryption value of P_i (or the intermediate decryption value of C_i) between E_0 and E_1 by X_i, i.e., $X_i = E_0(P_i) = E_1^{-1}(C_i)$. If (P_1, P_2) is a right pair with respect to the first differential, then $X_1 \oplus X_2 = \beta$. If both pairs (C_1, C_3) and (C_2, C_4) are right pairs with respect to the second differential, then $X_1 \oplus X_3 = \gamma = X_2 \oplus X_4$. If all these conditions are satisfied then $X_3 \oplus X_4 = \beta$. The boomerang attack uses the obtained β value by decrypting the pair (X_3, X_4), which with probability p_0 leads to $P_3 \oplus P_4 = \alpha$. The overall probability of such a quartet is $p_0^2 p_1^2$.

The attack can be mounted for all possible β's and γ's simultaneously (as long as $\beta \neq \gamma$). Thus, a right quartet for E is encountered with probability no less than $(\hat{p}_0 \hat{p}_1)^2$, where:

$$\hat{p}_0 = \sqrt{\sum_{\beta} \mathrm{Pr}^2[\alpha \to \beta]}, \qquad \text{and} \qquad \hat{p}_1 = \sqrt{\sum_{\gamma} \mathrm{Pr}^2[\gamma \to \delta]}.$$

The complete analysis is given in [36]. In particular it is possible to show that for a specific value of β, and the corresponding probability p_0 and all γ's simultaneously, the probability for $X_3 \oplus X_4 = \beta$ is $p_0 \hat{p_1}^2$. We shall use this fact later.

5.2 Differential-Bilinear-Boomerang Attack (and Relatives)

We first note that linear, differential-linear, and bilinear approximations, are special cases of differential-bilinear approximations (up to whether we consider pairs of plaintexts or plaintext/ciphertext pairs). Hence, if we can combine the differential-bilinear attack with some other attack, we can actually combine any of the linear, the differential-linear, or the bilinear attacks as well.

Our newly proposed attacks exploit the β difference between the intermediate decryption values X_3 and X_4 of the encryptions whose ciphertexts are C_3 and C_4. If there is a differential-bilinear approximation for E_0^{-1} (the decryption through

E_0), then the pair (X_3, X_4) has the required input difference, and thus, there is some bilinear relation between X_3 and X_4 whose probability (or bias) is non-trivial.

More formally, let (X_3, X_4) (generated by the partial decryption of C_3 and C_4 during the boomerang process) be with difference β. Assume that there exists a differential-bilinear approximation with bias $2pq^2$ for E_0^{-1} with input difference β. Thus, it is possible to analyze the corresponding plaintexts as in the differential-bilinear attack, just like as suggested in Section 3.

However, the pair (X_3, X_4) does not always have the required difference β, which occurs with probability $p_0\hat{p}_1^2$. By performing the analysis of the differential-bilinear attack again, and taking into consideration the probability that the β difference occurs, we conclude that the differential-bilinear relation has a bias of $2\hat{p}_1^2 p_0 pq^2$.

Actually, we treat the first sub-cipher E_0 as a cascade of two sub-sub-ciphers, i.e, $E_0 = E_{0_1} \circ E_{0_0}$. The differential is used in the the first part of the backward direction, i.e., in $E_{0_1}^{-1}$, while the bilinear approximation is used in the second par of $E_{0_0}^{-1}$ (also in the backward direction).

The differential-bilinear boomerang attack tries to obtain a difference between two intermediate encryption values in the transition between the first sub-sub-cipher and the second sub-sub-cipher (both are parts of the first sub-cipher). This is a somewhat “asymmetric” boomerang, where for the first pair (P_1, P_2) we have a different number of rounds in the first sub-cipher than for the pair (P_3, P_4).

As the bias of the differential-bilinear boomerang is very low, it might seem that using other techniques based on decomposing the cipher into sub-cipher is always better than this attack. Even though currently we have no example where this attack is better than other combinations, we believe such cases exist.

We start with showing that there are cases where the proposed attack can be better than the boomerang attack. At a first glance, even if we assume that the bias of the differential-bilinear approximation of E_0 is $1/2$, then the bias of the whole differential-bilinear boomerang approximation is $\hat{p}_1^2 p_0$. Thus, the data complexity of the differential-bilinear boomerang attack is expected to be at least $O(\hat{p}_1^{-4} p_0^{-2})$, while a regular boomerang attack requires a usually smaller data complexity of $O(\hat{p}_0^{-2}\hat{p}_1^{-2})$. However, this is true only for a boomerang attack that uses regular differentials. In such case, the probability of the differential in the decryption direction is equal to the probability in the encryption direction. But in some boomerang attacks, truncated differential are used, and for these kind of differentials the probability depends on the direction. Thus, it might lead to an attack which is better than the boomerang attack, if for example, there is a truncated differential that is used in the forward direction of E_0, but cannot be used in the backward direction due to low probability.

Another attack that can be used instead of the differential-bilinear boomerang is the differential-(bi)linear attack. As mentioned before, there is a good differential in the backward direction, and a good bilinear approximation. The reason why this process might yield a better attack is that the difference predicted by the differential after the partial decryption may not be suitable for concatenation with a bilinear approximation. In this case, the boomerang process is used to change the difference to a more “friendly” one.

For linear (or differential-linear) cryptanalysis, where the exact difference has a much smaller effect, the answer is different. Usually, it is assumed that the approximation has an independent random behavior for any two plaintexts, even if there is some constant difference between them. The chosen ciphertext linear cryptanalysis [23] has shown that this is not the case, and that the actual values encrypted can alter the probabilities related to the approximation. Hence, the bias of the linear approximation may increase if there is a specific difference, instead of some random difference. Such an increase would lead to an higher biases, which in turn would mean better attacks.

6 Summary

In this paper we presented several new combined attacks. Each of these combinations has scenarios where it yields an attack that may be better than differential-linear attacks, differential attacks, or linear attacks for some ciphers.

The differential-bilinear attack, the higher-order differential-linear attack, and the (differential-)(bi)linear boomerang attack, are examples of attacks based on treating the cipher as a cascade of sub-ciphers. This kind of treatment allows us to present a a differential-bilinear approximation for 6-round s^{5}DES with a bias of 1/8. The decomposition into sub-ciphers can be used to enlarge the linear weak-key class of IDEA by a factor of 512.

We conclude that new designs have to take into consideration combined attacks, including the well-known ones such as differential-linear and boomerang attacks, as well as the new ones presented in this paper.

Acknowledgments

We would like to thank Nicolas Courtois for the information he provided to us, that helped us a lot during the work on this paper, and the anonymous referees for suggesting valuable ideas and improvements.

References

1. Ross Anderson, Eli Biham, Lars R. Knudsen, *Serpent: A Proposal for the Advanced Encryption Standard*, NIST AES Proposal, 1998.
2. Eli Biham, *Higher Order Differential Cryptanalysis*, unpublished paper, 1994.
3. Eli Biham, *On Matsui’s Linear Cryptanalysis*, Advances in Cryptology, proceedings of EUROCRYPT ’94, Lecture Notes in Computer Science 950, pp. 341–355, Springer-Verlag, 1994.
4. Eli Biham, *Cryptanalysis of Ladder-DES*, proceedings of Fast Software Encryption 4, Lecture Notes in Computer Science 1267, pp. 134–138, Springer-Verlag, 1997.

5. Eli Biham, Adi Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.
6. Eli Biham, Alex Biryukov, Adi Shamir, *Miss in the Middle Attacks on IDEA and Khufu*, proceedings of Fast Software Encryption 6, Lecture Notes in Computer Science 1636, pp. 124–138, Springer-Verlag, 1999.
7. Eli Biham, Alex Biryukov, Adi Shamir, *Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials*, Advances in Cryptology, proceedings of EUROCRYPT '99, Lecture Notes in Computer Science 1592, pp. 12–23, Springer-Verlag, 1999.
8. Eli Biham, Orr Dunkelman, Nathan Keller, *Enhanced Differential-Linear Cryptanalysis*, Advances in Cryptology, proceedings of ASIACRYPT '02, Lecture Notes in Computer Science 2501, pp. 254–266, Springer-Verlag, 2002.
9. Eli Biham, Orr Dunkelman, Nathan Keller, *Differential-Linear Cryptanalysis of Serpent*, proceedings of Fast Software Encryption 10, Lecture Notes in Computer Science 2887, pp. 9–21, Springer-Verlag, 2003.
10. Alex Biryukov, Eyal Kushilevitz, *From Differential Cryptoanalysis to Ciphertext-Only Attacks*, Advances in Cryptology, proceedings of CRYPTO '98, Lecture Notes in Computer Science 1462, pp. 72–88, Springer-Verlag, 1998.
11. Alex Biryukov, Jorge Nakahara, Bart Preneel, Joos Vandewalle, *New Weak-Key Classes of IDEA*, proceedings of ICICS '02, Lecture Notes in Computer Science 2513, pp. 315–326, Springer-Verlag, 2002.
12. Alex Biryukov, Adi Shamir, *Structural Cryptanalysis of SASAS*, Advances in Cryptology, proceedings of EUROCRYPT '01, Lecture Notes in Computer Science 2045, pp. 394–405, Springer-Verlag, 2001.
13. Johan Borst, Lars R. Knudsen, Vincent Rijmen, *Two Attacks on Reduced Round IDEA*, Advances in Cryptology, proceedings of EUROCRYPT '97, Lecture Notes in Computer Science 1233, pp. 1–13, Springer-Verlag, 1997.
14. Nicolas T. Courtois, *Feistel Schemes and Bi-Linear Cryptanalysis*, Advances in Cryptology, proceedings of CRYPTO '04, Lecture Notes in Computer Science 3152, pp. 23–40, Springer-Verlag, 2004.
15. Nicolas T. Courtois, *Feistel Schemes and Bi-Linear Cryptanalysis (extended version)*, private communications, 2004.
16. Nicolas T. Courtois, *The Inverse S-box, Non-linear Polynomial Relations and Cryptanalysis of Block Ciphers*, private communications, 2004.
17. Joan Daemen, René Govaerts, Joos Vandewalle, *Weak Keys for IDEA*, Advances in Cryptology, proceedings of CRYPTO '93, Lecture Notes in Computer Science 773, pp. 224–231, Springer-Verlag, 1994.
18. Joan Daemen, Lars R. Knudsen, Vincent Rijmen, *The Block Cipher Square*, proceedings of Fast Software Encryption 4, Lecture Notes in Computer Science 1267, pp. 149–165, Springer-Verlag, 1997.
19. Niels Ferguson, John Kelsey, Stefan Lucks, Bruce Schneier, Mike Stay, David Wagner, Doug Whiting, *Improved Cryptanalysis of Rijndael*, proceedings of Fast Software Encryption 7, Lecture Notes in Computer Science 1978, pp. 213–230, Springer-Verlag, 2001.
20. Philip Hawkes, *Differential-Linear Weak Keys Classes of IDEA*, Advances in Cryptology, proceedings if EUROCRYPT '98, Lecture Notes in Computer Science 1403, pp. 112-126, Springer-Verlag, 1998.
21. Kwangjo Kim, Sangjun Lee, Sangjun Park, Daiki Lee, *How to Strengthen DES against Two Robust Attacks*, proceedings of Joint Workshop on Information Security and Cryptology, 1995.

22. Lars Knudsen, *Truncated and Higher Order Differentials*, proceedings of Fast Software Encryption 2, Lecture Notes in Computer Science 1008, pp. 196–211, Springer-Verlag, 1995.
23. Lars R. Knudsen, John E. Mathiassen, *A Chosen-Plaintext Linear Attack on DES*, proceedings of Fast Software Encryption 7, Lecture Notes in Computer Science 1978, pp. 262–272, Springer-Verlag, 2001.
24. Lars R. Knudsen, David Wagner, *Integral Cryptanalysis*, proceedings of Fast Software Encryption 9, Lecture Notes in Computer Science 2365, pp. 112–127, Springer-Verlag, 2002.
25. Xuejia Lai, *Higher Order Derivations and Differential Cryptanalysis*, in *Communications and Cryptography: Two Sides of One Tapestry*, Kluwer Academic Publishers, pp. 227-233, 1994.
26. Xuejia Lai, James L. Massey, *A Proposal for a New Block Cipher Encryption Standard*, Advances in Cryptology, proceedings of EUROCRYPT '90, Lecture Notes in Computer Science 473, pp. 389–404, Springer-Verlag, 1991.
27. Susan K. Langford, *Differential-Linear Cryptanalysis and Threshold Signatures*, Ph.D. thesis, 1995.
28. Susan K. Langford, Martin E. Hellman, *Differential-Linear Cryptanalysis*, Advances in Cryptology, proceedings of CRYPTO '94, Lecture Notes in Computer Science 839, pp. 17–25, Springer-Verlag, 1994.
29. Stefan Lucks, *The Saturation Attack — A Bait for Twofish*, proceedings of Fast Software Encryption 8, Lecture Notes in Computer Science 2355, pp. 1–15, Springer-Verlag, 2002.
30. Mitsuru Matsui, *Linear Cryptanalysis Method for DES Cipher*, Advances in Cryptology, proceedings of EUROCRYPT '93, Lecture Notes in Computer Science 765, pp. 386–397, Springer-Verlag, 1994.
31. Mitsuru Matsui, Atsuhiro Yamagishi, *A new method for known plaintext attack of FEAL cipher*, Advances in Cryptology, proceedings of EUROCRYPT '92, Lecture Notes in Computer Science 658, pp. 81–91, Springer-Verlag, 1993.
32. US National Bureau of Standards, *Data Encryption Standard*, Federal Information Processing Standards Publications No. 46, 1977.
33. Akihiro Shimizu, Shoji Miyaguchi, *Fast Data Encipherment Algorithm FEAL*, Advances in Cryptology, proceedings of EUROCRYPT '87, Lecture Notes in Computer Science 304, pp. 267–278, Springer-Verlag, 1988.
34. Yongsup Shin, Jongsung Kim, Guil Kim, Seokhie Hong, Sangjin Lee, *Differential-Linear Type Attacks on Reduced Rounds of SHACAL-2*, proceedings of ACISP 2004, Lecture Notes in Computer Science 3108, pp. 110–122, Springer-Verlag, 2004.
35. Serge Vaudenay, *Provable Security for Block Ciphers by Decorrelation*, Journal of Cryptology, Vol 16, Number 4, pp. 249–286, Springer-Verlag, 2003.
36. David Wagner, *The Boomerang Attack*, proceedings of Fast Software Encryption 6, Lecture Notes in Computer Science 1636, pp. 156–170, Springer-Verlag, 1999.