J. Cryptol. (2008) 21: 131–147
DOI: 10.1007/s00145-007-9013-7

Journal of
CRYPTOLOGY

© 2007 International Association for Cryptologic Research

On Probability of Success in Linear and Differential Cryptanalysis

Ali Aydın Selçuk

Department of Computer Engineering, Bilkent University, Ankara, 06800, Turkey
selcuk@cs.bilkent.edu.tr

Communicated by Eli Biham

Received 28 April 2003 and revised 7 August 2007
Online publication 14 September 2007

Abstract. Despite their widespread usage in block cipher security, linear and differential cryptanalysis still lack a robust treatment of their success probability, and the success chances of these attacks have commonly been estimated in a rather ad hoc fashion. In this paper, we present an analytical calculation of the success probability of linear and differential cryptanalytic attacks. The results apply to an extended sense of the term "success" where the correct key is found not necessarily as the highest-ranking candidate but within a set of high-ranking candidates. Experimental results show that the analysis provides accurate results in most cases, especially in linear cryptanalysis. In cases where the results are less accurate, as in certain cases of differential cryptanalysis, the results are useful to provide approximate estimates of the success probability and the necessary plaintext requirement. The analysis also reveals that the attacked key length in differential cryptanalysis is one of the factors that affect the success probability directly besides the signal-to-noise ratio and the available plaintext amount.

Key words. Block ciphers, Linear cryptanalysis, Differential cryptanalysis, Success probability, Order statistics.

1. Introduction

Differential cryptanalysis (DC) [1] and linear cryptanalysis (LC) [11,12] are two of the most important techniques in block cipher cryptanalysis today. Virtually every modern block cipher has its security checked against these attacks and a number of them have actually been broken. Despite this widespread utilization, evaluation of the success probability of these attacks is usually done in a rather ad hoc fashion: Success chances of differential attacks are typically evaluated according to the empirical observations based on the "signal-to-noise ratio" [1]. In the case of linear cryptanalysis, arbitrary ciphers are analyzed using Matsui's results for his DES attacks [11,12], which were in fact carefully calculated for and were specific to those particular attacks.

In this paper, we present an analytical, generally applicable calculation of the success probability of linear and differential attacks. Throughout the analysis, a generalized

definition of the term "success" is dealt with: If an attack on an m-bit key gets the correct value ranked among the top r out of 2^m possible candidates, we say the attack obtained an $(m - \lg r)$-bit *advantage* over exhaustive search. The traditional, more strict definition of success, where the attack discovers the right key as the first candidate, corresponds to obtaining an m-bit advantage over an m-bit key.

The analysis presented provides formulas for direct calculation of the success probability of linear and differential attacks. The amount of data required for an attack to achieve a certain success probability can also be calculated through these formulas. Furthermore, the analysis shows that the aimed advantage level—that is, in more traditional terms, the number of key bits attacked—is one of the factors that affect the success probability in differential cryptanalysis directly besides the already established factors of the signal-to-noise ratio and the expected number of right pairs.

The notations in the paper common to all sections include ϕ and Φ for the probability density and the cumulative distribution functions of the standard normal distribution. Besides, $\mathcal{B}$, $\mathcal{N}$, and $\mathcal{P}$ are used for denoting the binomial, normal, and Poisson distributions, respectively.

2. Success Probability in Linear Cryptanalysis

Linear cryptanalysis, developed by Matsui [11], is a known-plaintext attack that exploits the statistical correlation among the plaintext, ciphertext, and key bits of a block cipher to discover the encryption key. The first step in a linear attack is to find a *linear approximation* for the cipher. A linear approximation is a binary equation of the bits of the plaintext, ciphertext, and the key, which holds with a probability $p \neq 1/2$. The quantity $|p - 1/2|$, known as the *bias*, is a measure of correlation among the plaintext, ciphertext, and key bits, and it can be used to distinguish the actual key from random key values. In an attack, the attacker collects a large number of plaintext-ciphertext blocks, and for each possible key value he counts the number of plaintext-ciphertext blocks that satisfy the approximation. Assuming that the bias of the approximation with the right key will be significantly higher than the bias with a random key, the key value that maximizes the bias over the given plaintext sample is taken as the right key.

In general, it may be sufficient to have the right key ranked reasonably high among the candidates rather than having it as the absolute highest. For example, in Matsui's attack on DES, a 26-bit portion of the key was attacked where the right key was ranked among the top 2^{13}. In this kind of ranking attacks, all candidates ranked higher than the right key must be tried before the right key can be reached. Each candidate must be checked with all combinations of the remaining, unattacked bits to see if it is the right value. In such an attack, where an m-bit key is attacked and the right key is ranked rth among all 2^m candidates, the attack provides a complexity reduction by a factor of $2^{m-\lg r}$ over the exhaustive search. In our analysis, we refer to $m - \lg r$ as the *advantage* provided by the attack.

2.1. *Problem Statement*

Consider the problem where an attacker is interested in getting the right key ranked within the r top candidates among a total of 2^m keys, where an m-bit key is attacked,

with an approximation of probability p, using N plaintext blocks. Let k_0 denote the right key and k_i, $1 \le i \le 2^m - 1$, be the wrong key values, and let n denote $2^m - 1$. Let $X_i = T_i/N - 1/2$ and $Y_i = |X_i|$, where T_i is the counter for the plaintexts satisfying the approximation with key k_i. Let W_i, $1 \le i \le 2^m - 1$, be the Y_i, $i \ne 0$, sorted in increasing order. That is, W_1 is the lowest sample bias $|T_i/N - 1/2|$ obtained among the wrong keys, W_n is the highest. Then, the two conditions for the success of the attack are

$$X_0/(p - 1/2) > 0, \tag{1}$$

that is, $T_0/N - 1/2$ and $p - 1/2$ have the same sign, and

$$|X_0| > W_{n-r+1}. \tag{2}$$

In the rest of this analysis, we assume for simplicity that $p > 1/2$.[1] Hence, the two conditions become

$$X_0 > 0, \tag{3}$$

$$X_0 > W_{n-r+1}. \tag{4}$$

This modeling of the success probability was originally given by Junod [7], where he derived an expression of the success probability in terms of Euler's incomplete beta integral assuming that the T_is are independent and they are identically distributed for $i \ne 0$. He also presented a numerical calculation of that expression for Matsui's 26-bit DES attack [12] assuming that the approximation has a zero bias for a wrong key, i.e., $E[T_i/N - 1/2] = 0$ for $i \ne 0$.

Here, we present a more general calculation of the success probability using the normal approximation for order statistics. Like Junod, we also assume the independence of the T_i counters and a zero bias for the wrong keys. Since the zero bias for the wrong keys is the ideal case for an attacker, the results can be seen as an upper bound for the actual success probability.

2.2. *Order Statistics*

In this section we give a brief review of order statistics, as treated in [13]. Theorem 1, the key for our analysis, states the normal approximation for the order statistics.

Definition 1. Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent, identically distributed random variables. Arrange the values of $\xi_1, \xi_2, \ldots, \xi_n$ in increasing order, resulting in $\xi_1^*, \xi_2^*, \ldots, \xi_n^*$. The statistic ξ_i^* is called the *i-th order statistic* of the sample $\xi_1, \xi_2, \ldots, \xi_n$.

Definition 2. For $0 < q < 1$, the *sample quantile of order q* is the $\lfloor qn \rfloor + 1$-th order statistic $\xi^*_{\lfloor qn \rfloor + 1}$.

Theorem 1. *Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent, identically distributed random variables, with an absolutely continuous distribution function $F(x)$. Suppose that the*

[1] The corresponding results for the case $p < 1/2$ can easily be obtained by substituting $-X_0$ for X_0.

density function $f(x) = F'(x)$ *is continuous and positive on the interval* $[a, b)$. *If* $0 < F(a) < q < F(b) < 1$, *and if* $i(n)$ *is a sequence of integers such that*

$$\lim_{n\to\infty} \sqrt{n}\left|\frac{i(n)}{n} - q\right| = 0,$$

further if ξ_i^* *denotes* i*-th order statistic of the sample* $\xi_1, \xi_2, \ldots, \xi_n$, *then* $\xi_{i(n)}^*$ *is in the limit normally distributed, i.e.,*

$$\lim_{n\to\infty} P\left(\frac{\xi_{i(n)}^* - \mu_q}{\sigma_q} < x\right) = \Phi(x),$$

where

$$\mu_q = F^{-1}(q),$$
$$\sigma_q = \frac{1}{f(\mu_q)}\sqrt{\frac{q(1-q)}{n}}.$$

Taking $i(n) = \lfloor qn \rfloor + 1$, the theorem states that the empirical sample quantile of order q of a sample of n elements is for sufficiently large n nearly normally distributed with expectation $\mu_q = F^{-1}(q)$ and standard deviation $\sigma_q = \frac{1}{f(\mu_q)}\sqrt{\frac{q(1-q)}{n}}$.

2.3. *Success Probability*

The sample bias of the right key, $X_0 = T_0/N - 1/2$, approximately follows a normal distribution $\mathcal{N}(\mu_0, \sigma_0^2)$ with $\mu_0 = p - 1/2$ and $\sigma_0^2 = 1/(4N)$. The absolute sample bias of wrong keys, $Y_i, i \neq 0$, follow a folded normal distribution (see Appendix 4) $\mathcal{FN}(\mu_w, \sigma_w^2)$ with $\mu_w = 0$, assuming a zero bias for wrong keys, and $\sigma_w^2 = 1/(4N)$. We use f_0, F_0 and f_w, F_w to denote the probability density and the cumulative distribution functions of X_0 and $Y_i, i \neq 0$, respectively.

In an a-bit advantage attack on an m-bit key, success is defined as

$$X_0 > 0, \tag{5}$$
$$X_0 > W_{\bar{r}}, \tag{6}$$

where $W_1, W_2, \ldots, W_{2^m-1}$ are the absolute sample bias of the wrong keys sorted in increasing order, and $\bar{r}$ denotes $2^m - 2^{m-a}$. According to Theorem 1, $W_{\bar{r}}$ approximately follows a normal distribution $\mathcal{N}(\mu_q, \sigma_q^2)$, which we denote by F_q, where

$$\mu_q = F_w^{-1}(1 - 2^{-a}) = \mu_w + \sigma_w \Phi^{-1}(1 - 2^{-a-1}),$$
$$\sigma_q = \frac{1}{f_w(\mu_q)} 2^{-\frac{m+a}{2}} = \frac{\sigma_w}{2\phi(\Phi^{-1}(1 - 2^{-a-1}))} 2^{-\frac{m+a}{2}},$$

since F_w is folded normal. Then the probability of success, P_S, is

$$P_S = \int_0^\infty \int_{-\infty}^{x} f_q(y)\, dy f_0(x)\, dx. \tag{7}$$

For $a, m \geq 8$, we have $\mu_q > 5\sigma_q$ and, therefore, the probability of $W_{\bar{r}} < 0$ is negligible. Hence, (5) and (6) can be combined as

$$X_0 > W_{\bar{r}}. \tag{8}$$

Since both X_0 and $W_{\bar{r}}$ follow a normal distribution, $X_0 - W_{\bar{r}}$ follows a normal distribution too, which we denote by F_J, with mean $\mu_0 - \mu_q$ and variance $\sigma_0^2 + \sigma_q^2$. Therefore,

$$P_S = P(X_0 - W_{\bar{r}} > 0) = \int_0^\infty f_J(x)\,dx = \int_{-\frac{\mu_0-\mu_q}{\sqrt{\sigma_0^2+\sigma_q^2}}}^\infty \phi(x)\,dx. \tag{9}$$

Table 1 gives a numeric calculation of (9) for certain values of a and m, with $N = 8|p - 1/2|^{-2}$ plaintext blocks.

In Table 1, it is interesting that P_S is almost independent of the key length m for a given a. Note that, for $8 \leq a \leq 48$, σ_q satisfies $10^{-6} \leq \sigma_q/\sigma_0 \leq 10^{-1}$. Especially when dealing with success probabilities of 80% or more, the effect of σ_q is negligible and we can assume $\sqrt{\sigma_0^2 + \sigma_q^2} \approx \sigma_0$. Then (9) becomes

$$P_S = \int_{-\frac{\mu_0-\mu_q}{\sigma_0}}^\infty \phi(x)\,dx = \int_{-2\sqrt{N}(|p-1/2|-F_w^{-1}(1-2^{-a}))}^\infty \phi(x)\,dx, \tag{10}$$

where the success probability is a function of the advantage level a, independent of the number of key bits attacked m.

In (10), F_w is the folded normal distribution $\mathcal{FN}(0, \sigma_w^2)$, and $F_w^{-1}(1 - 2^{-a}) = \sigma_w \Phi^{-1}(1 - 2^{-a-1})$ for $\sigma_w = 1/(2\sqrt{N})$, yielding the following main result:

Theorem 2. *Let P_S be the probability that a linear attack on an m-bit subkey, with a linear approximation of probability p, with N known plaintext blocks, delivers an a-bit or higher advantage. Assuming that the linear approximation's probability to hold is independent for each key tried and is equal to $1/2$ for all wrong keys, we have, for sufficiently large m and N,*

$$P_S = \Phi\left(2\sqrt{N}|p - 1/2| - \Phi^{-1}(1 - 2^{-a-1})\right). \tag{11}$$

A numerical calculation of (11) is shown in Table 2, where the success probability is given as a function of the aimed advantage level a and c_N, the amount of available plaintexts as a multiple of $|p - 1/2|^{-2}$ (i.e., $c_N = N/|p - 1/2|^{-2}$). A comparison of the

Table 1. The success probability P_S according to (9) for obtaining an a-bit advantage on an m-bit key, for $N = 8|p - 1/2|^{-2}$ plaintexts.

a	$m = 8$	$m = 16$	$m = 32$	$m = 48$
8	0.996	0.997	0.997	0.997
16	–	0.903	0.909	0.909
32	–	–	0.250	0.248
48	–	–	–	0.014

Table 2. Probability of achieving an a-bit advantage with $N = c_N|p - 1/2|^{-2}$ plaintexts, according to (11).

a	$c_N = 2$	$c_N = 4$	$c_N = 8$	$c_N = 16$	$c_N = 32$	$c_N = 64$
8	0.477	0.867	0.997	1.000	1.000	1.000
16	0.067	0.373	0.909	1.000	1.000	1.000
32	0.000	0.010	0.248	0.952	1.000	1.000
48	0.000	0.000	0.014	0.552	0.999	1.000

columns of Table 1 to the column of Table 2 for $c_N = 8$ shows that the two are almost identical.

Equation (11) implies that $2\sqrt{N}|p - 1/2| - \Phi^{-1}(1 - 2^{-a-1}) = \Phi^{-1}(P_S)$, yielding a direct formula to calculate the required number of plaintexts to achieve a certain success probability P_S:

Corollary 1. *With the assumptions of Theorem 2,*

$$N = \left(\frac{\Phi^{-1}(P_S) + \Phi^{-1}(1 - 2^{-a-1})}{2}\right)^2 \cdot |p - 1/2|^{-2} \tag{12}$$

plaintext blocks are needed in a linear attack to accomplish an a-bit advantage with a success probability of P_S.

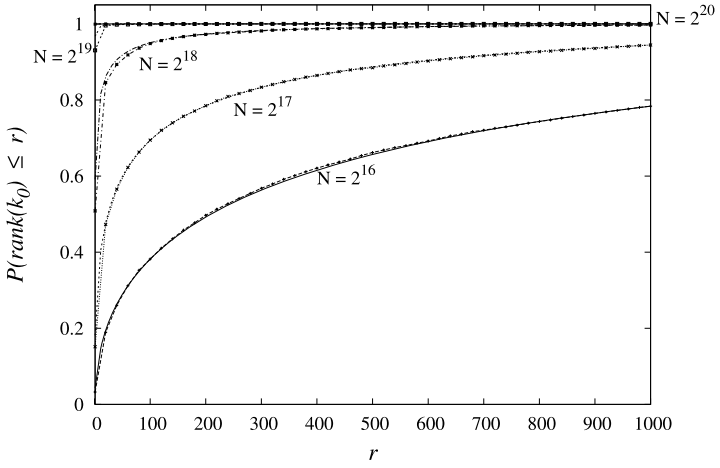
2.4. *Discussion of the Assumptions and Experimental Results*

In a typical linear attack, N is at least in the order of 2^{20} and p is very close to $1/2$. Hence, the normal approximation for the binomial T_i counters and for $X_i = (T_i/N - 1/2)$ can be expected to be extremely good.

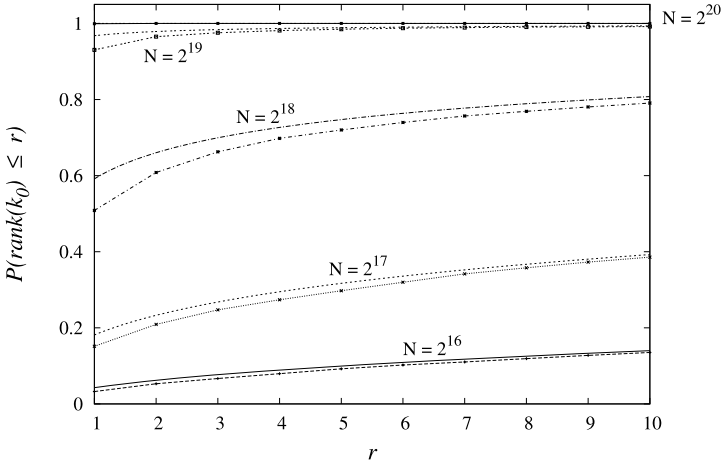
Regarding the normal approximation for the order statistics, the sample size is $n = 2^m - 1$, which can be expected to give accurate results for fairly large values of m [16]. For a practical evaluation, we implemented Matsui's 8-round DES attack [12] and compared the actual success probability to the results of (11). The attack uses a 6-round DES approximation with a bias $1.95 \cdot 2^{-9}$ and targets the keys of S5 in the first and the eighth rounds, with 12 key bits in total.[2] In the experiments the attack was run 10,000 times for each value of N.

The success probability according to the experimental results and according to (11) are compared in Fig. 1. Figure 1(a) shows that (11) gives a quite precise calculation of the success probability for most cases. The exception to this accuracy, as Figure 1(b) shows, is when the top ranking probability is of concern and with a relatively small P_S value, in which case about a 10% error rate is possible.

[2] The benefit of using DES in this experiment is that, it was observed in [15] that the bias of linear approximations of DES-like ciphers can be estimated accurately by the piling-up lemma [11], which is not always the case for other ciphers (e.g., RC5). Hence, using DES as the test cipher, the experiments can be conducted free of the errors that would result from a miscalculation of the bias.



(a) The results for the range $1 \leq \text{rank}(k_0) \leq 1000$. The theoretical and experimental results are mostly indistinguishable.



(b) The same plots with a focus on the range 1–10. Now a difference can be observed, especially when the top ranking probability is of concern. The theoretical and practical results are still indistinguishable for $N = 2^{20}$.

Fig. 1. A comparison of (11) with the experimental success rates. The bias of the linear approximation is $1.95 \cdot 2^{-9}$. The plots with the *linespoint* style show the experimental results; those with the *lines* style are P_S according to (11).

2.5. *Probability of Top Ranking*

A more precise calculation of the success probability is possible for the special case $a = m$ (i.e., when the right key is to be ranked the highest) which does not use the

Table 3. The top ranking probability $P(\text{rank}(k_0) = 1)$ in LC according to (11), (13), and the experimental results.

N	(11)	(13)	Exp.
2^{16}	0.043	0.038	0.033
2^{17}	0.181	0.159	0.151
2^{18}	0.592	0.539	0.509
2^{19}	0.968	0.949	0.930
2^{20}	0.999	0.999	0.999

normal approximation for order statistics. In this case, the success probability can be expressed directly as

$$\begin{aligned} P_S &= \int_0^{\infty} \left(\int_{-x}^{x} f_w(y)\, dy \right)^{2^m-1} f_0(x)\, dx \\ &= \int_{-2\sqrt{N}|p-1/2|}^{\infty} \left(\int_{-x-2\sqrt{N}|p-1/2|}^{x+2\sqrt{N}|p-1/2|} \phi(y)\, dy \right)^{2^m-1} \phi(x)\, dx, \end{aligned} \tag{13}$$

again assuming that the T_i counters are independent. A numerical comparison of (11), (13), and the experimental results is given in Table 3.

3. Success Probability in Differential Cryptanalysis

Differential cryptanalysis, developed by Biham and Shamir [1], is a chosen-plaintext attack that exploits the correlation between the input and output differences of a pair of plaintext blocks encrypted under the same key. The first step in a differential attack is to find a *characteristic* of the cipher attacked. A characteristic is a sequence of differences between the round inputs in the encryption of two plaintext blocks with a given initial difference. For a characteristic to be useful in an attack, a plaintext pair with the given initial difference must have a non-trivial probability to follow the given sequence of differences during encryption. Having obtained such a characteristic, the attacker collects a large number of plaintext-ciphertext pairs with the given initial difference. Assuming that the characteristic is followed at the inner rounds of the cipher, each pair will suggest a set of candidates for the last round key. When a pair is a "right pair", which followed the characteristic, the actual key will always be among the keys suggested. If the pair is "wrong", it may be detected and discarded, or, otherwise, it will suggest a set of random keys. After processing all collected pairs and counting the keys they suggest, the key value that is suggested most will be taken as the right key.

An important measure for the success of a differential attack is the proportion of the probability of the right key being suggested by a right pair to the probability of a random key being suggested by a random pair with the given initial difference. This proportion is called the "signal-to-noise ratio". Biham and Shamir [1] observed a strong relation between the signal-to-noise ratio and the success chance of an attack. By empirical evidence, they suggested that when the signal-to-noise ratio is around 1–2, about 20–40 right pairs would be sufficient; and when the signal-to-noise ratio is much higher, 3–4 right pairs would usually be enough.

3.1. *Distribution Parameters*

We use a notation similar to the one used for linear cryptanalysis: m is the number of key bits attacked; N denotes the total number of pairs analyzed. k_0 denotes the right key, $k_i, 1 \le i \le 2^m - 1$, denote the wrong keys. p_i is the probability of k_i being suggested by a plaintext pair; T_i counts the number of times k_i is suggested. $W_i, 1 \le i \le 2^m - 1$, denote $T_i, i \neq 0$, sorted in increasing order. The probability of the characteristic is denoted by p, and $\mu = pN$ denotes the expected number of right pairs. p_r is the average probability of some given key being suggested by a random pair with the given initial difference. S_N denotes the signal-to-noise ratio, p/p_r.

In our analysis, we assume that the T_i values are independent and that they are identically distributed for $i \neq 0$. The latter assumption means that all wrong keys have the same chance of being suggested by a random pair. That is, all $p_i, i \neq 0$, are identical. We denote this probability by p_w.

The T_i counters have a binomial distribution, $\mathcal{B}(N, p_0)$ for T_0 and $\mathcal{B}(N, p_w)$ for $T_i, i \neq 0$. We denote these distribution functions by F_0 and F_w, and their density functions by f_0 and f_w, respectively. Typically, N is very large and therefore these binomial distributions can be approximated by normal distributions, $\mathcal{N}(\mu_0, \sigma_0^2)$ and $\mathcal{N}(\mu_w, \sigma_w^2)$, where the distribution parameters are,

$$\begin{aligned} \mu_0 &= p_0 N, & \sigma_0^2 &= p_0(1-p_0)N \approx p_0 N, \\ \mu_w &= p_w N, & \sigma_w^2 &= p_w(1-p_w)N \approx p_w N. \end{aligned}$$

In a typical differential attack, the right key gets counted by right pairs and also gets random hits from wrong pairs, while the wrong keys only gets hits from wrong pairs. In this case, we have

$$\begin{aligned} p_0 &= p + (1-p)p_r \approx p + p_r, \\ p_w &= p_r. \end{aligned}$$

Note that this typical behavior does not always happen and in certain exceptional cases (e.g., [5]) the right key gets counted only by right pairs without getting any random hits from wrong pairs. In such attacks, we have $p_0 = p$. The following analysis assumes the more typical case where $p_0 = p + p_r$ but can easily be extended to cases where $p_0 = p$ by substituting p for p_0 in (16) and doing the following derivations accordingly.

3.2. *Success Probability*

In an a-bit advantage attack, success is defined by getting k_0 ranked within the top 2^{m-a} candidates; that is, $T_0 > W_{2^m - 2^{m-a}}$. We denote $2^m - 2^{m-a}$ by $\bar{r}$.

An analysis along the same lines as the one on linear cryptanalysis—with the only major difference being that the T_is here have a normal distribution whereas the Y_is in linear cryptanalysis had a folded normal—gives

$$P_S = \int_{-\frac{\mu_0 - \mu_q}{\sqrt{\sigma_0^2 + \sigma_q^2}}}^{\infty} \phi(x)\, dx, \tag{14}$$

where $\mu_q = \mu_w + \sigma_w \Phi^{-1}(1-2^{-a})$, $\sigma_q = \frac{\sigma_w}{\phi(\Phi^{-1}(1-2^{-a}))} 2^{-\frac{m+a}{2}}$. For $\sigma_q^2 \ll \sigma_0^2$, we have

$$P_S = \int_{-\frac{\mu_0-\mu_q}{\sigma_0}}^{\infty} \phi(x)\,dx. \tag{15}$$

The lower bound of the integral can be written in terms of the signal-to-noise ratio as,

$$\begin{aligned}
\frac{-\mu_0+\mu_q}{\sigma_0} &= \frac{-p_0 N + p_w N + \sqrt{p_w N}\Phi^{-1}(1-2^{-a})}{\sqrt{p_0 N}} \\
&= \frac{-pN + \sqrt{p_r N}\Phi^{-1}(1-2^{-a})}{\sqrt{(p+p_r)N}} \\
&= -\sqrt{pN}\sqrt{\frac{p}{p+p_r}} + \sqrt{\frac{p_r}{p+p_r}}\Phi^{-1}(1-2^{-a}) \\
&= -\sqrt{\mu}\sqrt{\frac{S_N}{S_N+1}} + \sqrt{\frac{1}{S_N+1}}\Phi^{-1}(1-2^{-a}).
\end{aligned} \tag{16}$$

Hence, the following result is obtained for the success probability:

Theorem 3. *Let P_S be the probability that a differential attack on an m-bit key, with a characteristic of probability p and signal-to-noise ratio S_N, and with N plaintext-ciphertext pairs, delivers an a-bit or higher advantage. Assuming that the key counters are independent and that they are identically distributed for all wrong keys, we have, for sufficiently large m and N, and μ denoting pN,*

$$P_S = \Phi\left(\frac{\sqrt{\mu S_N} - \Phi^{-1}(1-2^{-a})}{\sqrt{S_N+1}}\right). \tag{17}$$

A numerical calculation of (17) for $S_N = 1$ and $S_N = 1000$ is given in Table 4 to provide a comparison with Biham and Shamir's empirical results [1]. The values very much agree with their observations for large S_N. For small S_N, the suggested 20–40 right pairs give a good success rate only for $a < 20$. To have a good success rate for larger values of a as well, 80 or more right pairs may be needed.

Note that in (17), when S_N is very large, $(\sqrt{\mu S_N} - \Phi^{-1}(1-2^{-a}))/\sqrt{S_N+1} \approx \sqrt{\mu}$ and $P_S \approx \Phi(\sqrt{\mu})$. Hence, for large S_N, we can talk about the success probability as a function of μ only, independent of a, as it was discussed by Biham and Shamir [1].

As in linear cryptanalysis, we can use (17) to get a direct formulation of the required number of plaintext-ciphertext pairs to achieve a certain success probability P_S:

Corollary 2. *With the assumptions of Theorem 3,*

$$N = \frac{(\sqrt{S_N+1}\Phi^{-1}(P_S) + \Phi^{-1}(1-2^{-a}))^2}{S_N} p^{-1} \tag{18}$$

plaintext-ciphertext pairs are needed in a differential attack to accomplish an a-bit advantage with a success probability of P_S.

Table 4. Probability of achieving an a-bit advantage for various values of the expected number of right pairs μ, according to (17).

a	$\mu = 20$	$\mu = 40$	$\mu = 60$	$\mu = 80$	$\mu = 100$	$\mu = 120$
8	0.900	0.995	1.000	1.000	1.000	1.000
16	0.585	0.936	0.994	1.000	1.000	1.000
32	0.107	0.527	0.858	0.973	0.996	1.000
48	0.010	0.151	0.490	0.794	0.942	0.988

(a) $S_N = 100$

a	$\mu = 4$	$\mu = 5$	$\mu = 6$	$\mu = 7$	$\mu = 8$	$\mu = 9$
8	0.972	0.984	0.991	0.995	0.997	0.998
16	0.969	0.982	0.990	0.994	0.997	0.998
32	0.964	0.979	0.988	0.993	0.996	0.998
48	0.960	0.977	0.986	0.992	0.995	0.997

(b) $S_N = 1000$

3.2.1. *Alternative Models*

It can reasonably be said that the normal approximation for the binomial counters may not be accurate when $Np(1-p) < 4$; hence alternative distributions should be considered in the attacks where S_N is high and the number of rights pairs is accordingly low. For such attacks, Poisson approximations $\mathcal{P}(\mu_0)$ and $\mathcal{P}(\mu_w)$ or the original binomial distributions $\mathcal{B}(N, p_0)$ and $\mathcal{B}(N, p_w)$ can be preferred to model the counters.

Although more accurate results are possible by the Poisson or binomial distributions, these distributions lack the main advantage of using the normal distribution, which is to provide closed form expressions such as (17) and (18). Besides, as the results in Section 3.4 show, there does not appear to be a significant difference between the results obtained by the normal approximation and those obtained by the original binomial distribution, even for smaller values of μ_0. Hence, the normal approximation appears to be mostly satisfactory. The experimental results point out a different, more inherent limitation however, which is discussed in Section 3.4.

3.3. *Probability of Top Ranking*

Similar to that in LC, a more precise, direct calculation of the success probability of a differential attack is possible for the special case $a = m$ (i.e., when the right key is to be ranked the highest) which does not use the normal approximation for order statistics:

$$\begin{aligned} P_S(m) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{x} f_w(y)\, dy \right)^{2^m-1} f_0(x)\, dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{x\sqrt{S_N+1}+\sqrt{\mu S_N}} \phi(y)\, dy \right)^{2^m-1} \phi(x)\, dx, \end{aligned} \tag{19}$$

again assuming the independence of the counters and the normal approximation for the binomial distribution.

3.4. *Discussion of the Assumptions and Experimental Results*

There were three main assumptions employed in developing the results in this section:

1. The binomial T_i counters can be approximated by the normal distribution.
2. The T_i counters can be taken to be independent.
3. Order statistics can be approximated by the normal distribution.

The normal approximation to the order statistics can be expected to behave similar to that in LC. However, the other two assumptions appear to be less accurate in DC than in LC:

First of all, the binomial counters $T_i \sim \mathcal{B}(N, p_i)$, may not be accurately modeled by the normal distribution unless $Np_i(1-p_i) \geq 4$. This may not be so much of a problem for T_0, since in a differential attack typically 4 or more expected right pairs are used. But for $T_i, i \neq 0$, the same may not hold true. Especially in attacks with a large S_N ratio, $Np_i(1-p_i), i \neq 0$, will be much less than 4.

Regarding the assumption of independent counters, in a differential attack, every plaintext-ciphertext pair suggests on average a certain number of key candidates. For instance, in a DES attack, on average 4 keys are suggested per s-box by a plaintext-ciphertext pair. Consequently, the key counters T_i in a differential attack sum up to a certain value, and hence are inherently correlated.

For a practical evaluation of these assumptions, we tested the derived equations in the 6-round DES attack of Biham and Shamir [1]. This attack uses a 3-round characteristic with $p = 1/16$, $S_N = 2^{16}$, and aims to discover 30 bits of the 6th round key, namely the keys of the s-boxes 2, 5, 6, 7, 8. To also test the equations in an attack with a low S_N ratio, we took a variant of this attack with the same characteristic where only the key of S5 is attacked. In this variant, we have $m = 6$, $p = 1/8$, and $S_N = 2$. The attacks were run 10,000 times for each value of μ. The results are summarized in Fig. 2.

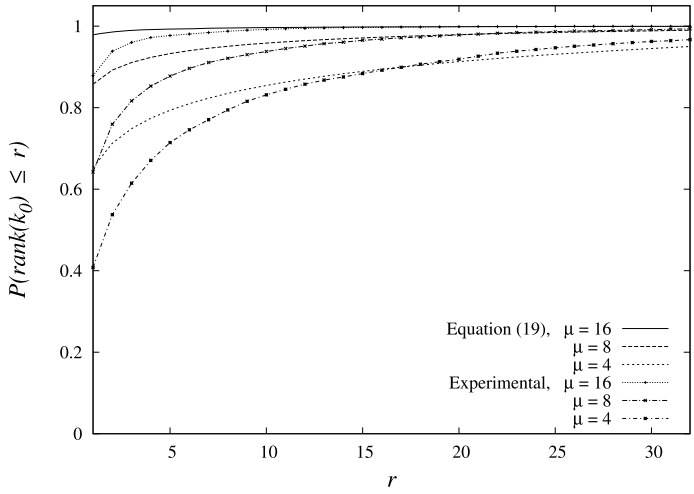
Figure 2 shows that (17) is not as accurate to calculate the success probability as its counterpart in LC. Nevertheless, when a success probability of 99% or higher is of interest, (17) gives a quite reliable estimate for P_S. For lower values, the results obtained through (17) may have a 30% or higher error rate.

To trace the source of this error, we first looked into the normal approximation for the binomial distribution. Recall that in formulation of the success probability,

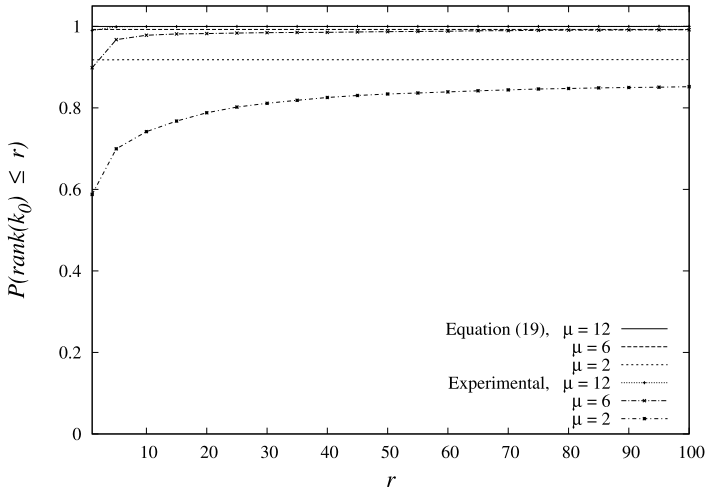
$$P_S = P(T_0 > W_{\bar{r}}) = \int_0^{\infty} \int_{-\infty}^{x} f_q(y)\, dy f_0(x)\, dx. \tag{20}$$

Equation (17) was derived from (20) assuming the normal distribution for f_0, $\mathcal{N}(\mu_0, \sigma_0^2)$. Also, $\mu_q = F_w^{-1}(q)$ and $\sigma_q = \frac{1}{f_w(\mu_q)}\sqrt{\frac{q(1-q)}{n}}$ were calculated assuming f_w was $\mathcal{N}(\mu_w, \sigma_w^2)$. In Fig. 3, we instead calculated (20) using $f_0 = \mathcal{B}(N, p_0)$, $f_w = \mathcal{B}(N, p_w)$ without the normal approximation.[3] The plots show that the results obtained are not much better than those of (17) with the normal approximation.

[3] Poisson approximations $\mathcal{P}(\mu_0)$ and $\mathcal{P}(\mu_w)$ can also be used here for the binomial f_0 and f_w if a more efficient calculation is desired. The Poisson approximation yielded very similar results to those obtained with the binomial distribution presented here.



(a) Attack on s-box 5; $S_N = 2$.

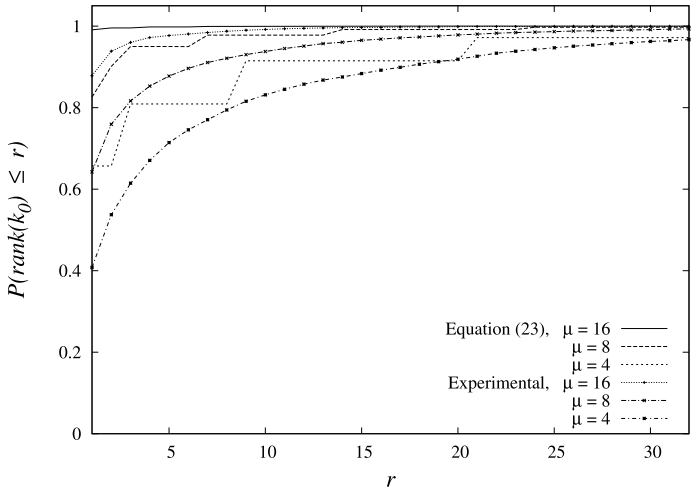


(b) Attack on s-boxes 2, 5, 6, 7, 8; $S_N = 2^{16}$.

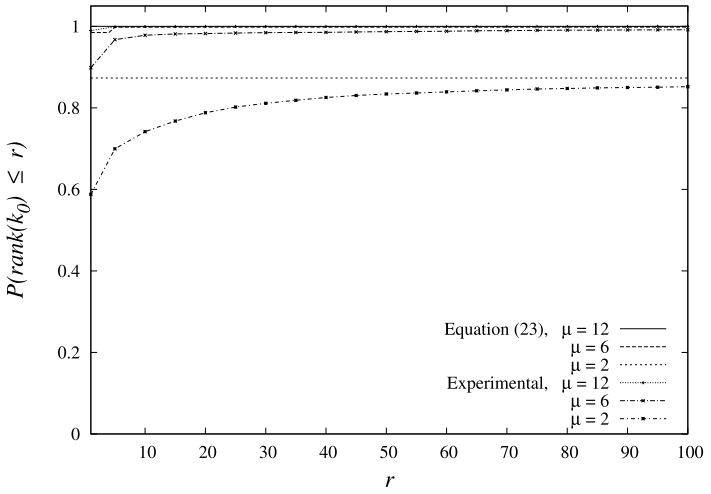
Fig. 2. A comparison of (17) and the experimental success rates of the 6-round DES attacks tested. The results show a considerable difference for the lower values of μ.

As another source of error, we turned to the normal approximation for the order statistics and, for a comparison, calculated the top ranking probabilities according to (19) which does not make use of this approximation, with $f_0 = \mathcal{B}(N, p_0)$, $f_w = \mathcal{B}(N, p_w)$. The results are shown in Table 5.

Results of (19), although somewhat better, turn out not to be much more accurate than those of (17), with a possible error rate as high as 30%. Equation (19) was derived without the normal approximation for the order statistics nor for the binomial distribution, the only major assumption employed being that the key counters T_i can be taken



(a) Attack on s-box 5; $S_N = 2$.



(b) Attack on s-boxes 2, 5, 6, 7, 8; $S_N = 2^{16}$.

Fig. 3. A comparison of (20) and the experimental success rates of the DES attacks tested. The binomial distribution does not provide any significant improvements over the normal approximation. The step-like behavior of (20) is due to the discrete nature of the binomial distribution used for F_0 and F_w.

to be independent. Hence, it appears that neglecting the dependence of the counters in DC is causing a non-negligible error in the success probability calculation.

This analysis shows that the dependence among the key counters is the principal source of the error observed in this section. As mentioned earlier, each plaintext-ciphertext pair suggests a certain number of keys in DC, and hence the key counters are inherently correlated. The results demonstrate that treating these counters independently can be a significant source of error in success probability calculations.

Table 5. $P(\text{rank}(k_0) = 1)$ according to (17), (20), (19), and the experimental results.

μ	(17)	(20)	(19)	Exp.
1	0.336	0.318	0.314	0.228
2	0.466	0.476	0.422	0.291
4	0.653	0.657	0.613	0.408
8	0.858	0.826	0.847	0.643
16	0.979	0.991	0.981	0.878
32	0.999	0.999	0.999	0.988
64	1.000	1.000	1.000	1.000

(a) $S_N = 2$

μ	(17)	(20)	(19)	Exp.
1	0.836	0.644	0.602	0.449
2	0.918	0.873	0.769	0.588
4	0.976	0.915	0.925	0.800
6	0.992	0.985	0.985	0.899
8	0.997	0.998	0.998	0.952
12	0.999	0.999	0.999	0.990
16	1.000	1.000	1.000	0.998

(b) $S_N = 2^{16}$

Note that assuming the key counters to be independent random variables is a very fundamental assumption for any general analysis of the success probability and, therefore, these results point out what appears to be a fundamental limitation of analytical success probability calculations for DC.

On the positive side, the equations derived in this section—(17), as well as (18), (19), and (20)—can be used reliably as long as the success probability of interest is 99% or higher. The equations can be useful for the lower values of the success probability as well, where they can be used to obtain rough estimates for P_S or N.

4. Conclusions

In this paper, we gave a formal probabilistic model of success in linear and differential cryptanalysis. We also provided efficient formulations that can be used to estimate the success probability of a given attack or to find its plaintext requirement to achieve a certain success level.

Experimental results show that the formulas developed for LC are quite precise, especially when a success probability of 90% or higher is of interest. The formulas appear to be less accurate for DC. The fact that the key counters are inherently correlated constitutes a fundamental difficulty for a simple and general formulation. Nevertheless, the equations derived neglecting this correlation turn out to provide reasonably accurate estimates for the higher values of the success probability. For the lower values, the equations can still be useful to obtain a rough estimate for the success probability or for the plaintext requirement.

It must be noted that, in the analysis for LC, it was assumed that the linear approximation would have a negligible bias for a wrong key. This assumption may not be true for some ciphers (e.g., for RC5 [14]), in which case the results obtained for the success probability here must be seen as an upper bound rather than an exact estimate, since having a zero bias for the wrong keys constitutes the ideal case for the attacker. On a separate note, our notion of "advantage" does not include the one bit of key information derived in a linear attack from the exclusive-or of the key bits on the right-hand side of the linear approximation. Counting that bit of information—if the bits included in the exclusive-or are not all included among the key bits derived—the advantage of the attack can be seen as $a + 1$.

There are several significant open problems in analyzing the success probability of cryptanalytic attacks. Finding a more accurate formulation of the success probability in

DC than those discussed in this paper would be a significant contribution. Success probability of different kinds of attacks such as differential-linear cryptanalysis [6], linear cryptanalysis with multiple approximations [4,9], boomerang attacks [17], or attacks with impossible differentials [2,3] can also be analyzed. On a more general theme, in this paper we discussed the success probability of a simple ranking attack, where the key portion attacked is attacked as a single part using a single ranking procedure. Analyzing the success probability of compound ranking attacks, where parts of the key are derived in separate attacks and then combined (e.g., Matsui's attack on the 16-round DES [12]), would be an important contribution. In particular, the success probability can be studied according to the optimal key ranking procedure of Junod and Vaudenay [8] for combining independently attacked key bits using a Neyman-Pearson approach.

Acknowledgements

I would like to thank to Ali Bıçak, Pascal Junod, and Burgess Davis for helpful discussions and comments, and to Murat Ak, Kamer Kaya, and Zahir Tezcan for their support in the implementation of the experiments. I am also grateful to Eli Biham and anonymous *J. Cryptology* referees whose comments and suggestions helped greatly to improve the paper.

Appendix The Folded Normal Distribution

When a normal random variable is taken without its algebraic sign, the negative side of the probability density function becomes geometrically folded onto the positive side. That is, if X has a normal distribution $\mathcal{N}(\mu, \sigma^2)$ with density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$$

then $Y = |X|$ has the density function

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}}\left(e^{-\frac{(y-\mu)^2}{2\sigma^2}} + e^{-\frac{(y+\mu)^2}{2\sigma^2}}\right), \quad y \geq 0.$$

The distribution of Y is called a *folded normal distribution* [10], which we denote by $\mathcal{FN}(\mu, \sigma^2)$. The mean and variance of Y are,

$$E(Y) = \mu(1 - 2\Phi(-\mu/\sigma)) + 2\sigma\phi(\mu/\sigma),$$
$$\mathrm{Var}(Y) = \mu^2 + \sigma^2 - E(Y)^2.$$

References

[1] E. Biham, A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard* (Springer, Berlin, 1993)

[2] E. Biham, A. Biryukov, A. Shamir, Cryptanalysis of skipjack reduced to 31 rounds using impossible differentials, in *Advances in Cryptology—Eurocrypt'99*, ed. by J. Stern. LNCS, vol. 1592 (Springer, Berlin, 1999), pp. 12–23
[3] E. Biham, A. Biryukov, A. Shamir, Miss in the middle attacks on IDEA, Khufu, and Khafre, in *Fast Software Encryption, 6th International Workshop*, ed. by L. Knudsen. LNCS, vol. 1636 (Springer, Berlin, 1999), pp. 124–138
[4] A. Biryukov, C. De Cannière, M. Quisquater, On multiple linear approximations, in *Advances in Cryptology—Crypto'04*. LNCS, vol. 3152 (Springer, Berlin, 2004), pp. 1–22
[5] L. Granboulan, Flaws in differential cryptanalysis of Skipjack, in *Fast Software Encryption, 8th International Workshop*, ed. by M. Matsui. LNCS, vol. 2355 (Springer, Berlin, 2001), pp. 328–336
[6] M. Hellman, S. Langford, Differential-linear cryptanalysis, in *Advances in Cryptology—Crypto'94*, ed. by Y.G. Desmedt. LNCS, vol. 839 (Springer, Berlin, 1994), pp. 17–25
[7] P. Junod, On the complexity of Matsui's attack, in *Selected Areas in Cryptography'01*. LNCS, vol. 2259 (Springer, Berlin, 2001), pp. 199–211
[8] P. Junod, S. Vaudenay, Optimal key ranking procedures in a statistical cryptanalysis, in *Fast Software Encryption, 10th International Workshop*. LNCS, vol. 2887 (Springer, Berlin, 2003), pp. 235–246
[9] S.B. Kaliski, M.J. Robshaw, Linear cryptanalysis using multiple approximations, in *Advances in Cryptology—Crypto'94*, ed. by Y.G. Desmedt. LNCS, vol. 839 (Springer, Berlin, 1994), pp. 26–39
[10] F.C. Leone, N.L. Nelson, R.B. Nottingham, The folded normal distribution, *Technometrics* **3**, 543–550 (1961)
[11] M. Matsui, Linear cryptanalysis method for DES cipher, in *Advances in Cryptology—Eurocrypt'93*, ed. by T. Helleseth. LNCS, vol. 765 (Springer, Berlin, 1993), pp. 386–397
[12] M. Matsui, The first experimental cryptanalysis of the Data Encryption Standard, in *Advances in Cryptology—Crypto'94*, ed. by Y.G. Desmedt. LNCS, vol. 839 (Springer, Berlin, 1994), pp. 1–11
[13] A. Rényi, *Probability Theory* (American Elsevier, New York, 1970)
[14] A.A. Selçuk, New results in linear cryptanalysis of RC5, in *Fast Software Encryption, 5th International Workshop*, ed. by S. Vaudenay. LNCS, vol. 1372 (Springer, Berlin, 1998), pp. 1–16
[15] A.A. Selçuk, On bias estimation in linear cryptanalysis, in *Indocrypt 2000*. LNCS, vol. 1977 (Springer, Berlin, 2000), pp. 52–66
[16] R.J. Serfling, *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics (Wiley, New York, 1980)
[17] D. Wagner, The boomerang attack, in *Fast Software Encryption, 6th International Workshop*, ed. by L. Knudsen. LNCS, vol. 1636 (Springer, Berlin, 1999), pp. 156–170