On Probability of Success in Differential and Linear Cryptanalysis

Ali Aydın Selçuk[1]

Network Systems Lab, Department of Computer Science, and
Center for Education and Research in Information Assurance and Security
Purdue Univeristy
West Lafayette, IN, 47907, USA
selcuk@cs.purdue.edu

CERIAS TR 2002-02

January 30, 2002

[1]Visiting scholar at the Network Systems Lab, supported by Professor Kihong Park.

Abstract

Differential and linear cryptanalysis, two of the most important techniques in modern block cipher cryptanalysis, still lack a sound, generally-applicable analysis of their success probabilities. In this paper, we present an analytical calculation of the success probability of differential and linear cryptanalytic attacks. In addition to a formulation of the success probability, the analysis yields a formulation of the attacks' data requirements as well, and it also finds the size of the aimed key information in differential cryptanalysis as one of the factors affecting the success probability. We also discuss the issue of key dependence in linear cryptanalysis, which can be a serious limitation for this technique of attack.

1 Introduction

Differential and linear cryptanalysis are the two 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 based on the empirical observations of Biham and Shamir [1] using the "signal-to-noise ratio". In the case of linear cryptanalysis, arbitrary ciphers are being analyzed by using the probability results of Matsui's DES attacks [7, 8], which were in fact calculated specifically for those attacks.

In this paper, we present a general analysis of the success probability in linear and differential cryptanalysis. We work with an extended definition of "success": If an attack on an m-bit key gets the correct value as the rth candidate among the 2^m possibilities, 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.

We present analytical calculations for the probability of success in linear and differential cryptanalysis for achieving a desired advantage level. The results also provide formulae for directly calculating the required amount of plaintext-ciphertext data for obtaining a given advantage with a given probability. In the case of differential cryptanalysis, the results show the aimed advantage level—that is, in more traditional terms, the number of key bits attacked—as a factor affecting the probability of success, in addition to the already established factors of the signal-to-noise ratio and the expected number of right pairs.

Before concluding, we briefly discuss the issue of non-negligible wrong key biases in linear cryptanalysis, which may be a limitation for this technique in certain settings.

Most notations are defined in the sections they are used. Notations common to all sections include ϕ and Φ for the probability density and the cumulative distribution functions of the standard normal distribution; $\mathcal{B}$ and $\mathcal{N}$ are used for denoting the binomial and normal distributions.

2 Success Probability in Linear Cryptanalysis

In a linear attack, the first step 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 $k_i, 1 \leq i \leq 2^m - 1$, 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 $W_i, 1 \leq i \leq 2^m - 1$, be the $Y_i, i \neq 0$, 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

$$\begin{aligned} X_0 &> 0, & (3)\\ X_0 &> W_{n-r+1}. & (4) \end{aligned}$$

This modeling of the success probability was originally given by Junod [4], 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 $i \neq 0$. He also presented a numerical calculation of that expression for Matsui's 26-bit DES attack [8] assuming that the approximation has a zero bias for a wrong key, i.e., $E[T_i/N - 1/2] = 0$ for $i \neq 0$.

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.

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

2.2 Order Statistics

In this section we give a brief review of order statistics, as treated in [9]. 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 $\xi_1^*, \xi_2^*, \ldots, \xi_n^*$. 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 $\xi^*_{\lfloor qn \rfloor}$.

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 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

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

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 $\mathcal{FN}(\mu_W, \sigma_W^2)$ (see Appendix A) 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

$$
\begin{aligned}
X_0 &> 0 && (5)\\
X_0 &> W_{\bar{r}} && (6)
\end{aligned}
$$

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

$$
\begin{aligned}
\mu_q &= F_w^{-1}(1-2^{-a}) = \mu_W + \sigma_W \Phi^{-1}(1-2^{-a-1})\\
\sigma_q &= \frac{1}{f_w(\mu_q)} 2^{-\frac{m+a}{2}} = \frac{\sigma_W}{2\phi(\Phi^{-1}(1-2^{-a-1}))} 2^{-\frac{m+a}{2}},
\end{aligned}
$$

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,

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

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

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 1: The success probability P_S according to equation (9) for obtaining an a-bit advantage on an m-bit key, for $N = 8|p-1/2|^{-2}$ plaintexts. It is interesting to note that P_S does not change much depending on m for a given a.

σ_q^2 is typically much smaller than σ_0^2. For $8 \le a \le 48$, we have $10^{-6} \le \sigma_q/\sigma_0 \le 10^{-1}$. Especially when dealing with success probabilities of 80% or more, the effect of σ_q is negligible. Assuming $\sqrt{\sigma_0^2 + \sigma_q^2} \approx \sigma_0$, (9) becomes

$$
\begin{aligned}
P_S &= \int_{-\frac{\mu_0-\mu_q}{\sigma_0}}^{\infty} \phi(x)\,dx && (10)\\
&= \int_{-2\sqrt{N}(|p-1/2|-F_w^{-1}(1-2^{-a}))}^{\infty} \phi(x)\,dx\,, && (11)
\end{aligned}
$$

independent of m, the number of key bits attacked. For F_w being the folded normal distribution $\mathcal{FN}(0,\sigma_W^2)$, we have $F_w^{-1}(1-2^{-a}) = \sigma_W \Phi^{-1}(1-2^{-a-1})$ and, for $\sigma_W = 1/(2\sqrt{N})$,

$$P_S = \int_{-2\sqrt{N}|p-1/2|+\Phi^{-1}(1-2^{-a-1})}^{\infty} \phi(x)\,dx\,. \qquad (12)$$

A numerical calculation of the success probability as expressed in (12) is given in Table 2.

Note that (10) is in fact the probability of $X_0 > E[W_{\bar{r}}]$, neglecting the variation in $W_{\bar{r}}$. A comparison of Table 1 and the column for $c_N = 8$ in Table 2 reveals that σ_q, the variance of $W_{\bar{r}}$, is quite insignificant and neglecting it is reasonable.

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

Table 2: Probability of achieving an a-bit advantage for various values of the plaintext amount $N = c_N|p-1/2|^{-2}$, according to equation (12).

The following theorem summarizes the main result of this section:

Theorem 2 *Let P_S be the probability that an Algorithm-2 linear attack as described in [7], where all candidates are tried for an m-bit subkey, in an approximation of probability p, with N known plaintext blocks, delivers an a-bit or higher advantage. Assuming that the approximation's probability is independent for each key tried and is equal to 1/2 for all wrong keys, we have, for sufficiently large m and N,*

$$P_S = \int_{-2\sqrt{N}|p-1/2|+\Phi^{-1}(1-2^{-a-1})}^{\infty} \phi(x)\,dx\,, \qquad (13)$$

independent of m.

Equation (13) implies $2\sqrt{N}|p-1/2| - \Phi^{-1}(1-2^{-a-1}) = \Phi^{-1}(P_S)$, yielding Corollary 1. This corollary gives a direct formula for the plaintext amount required for a desired success probability. The needed Φ^{-1} values can easily be calculated numerically, or they can be obtained from the standard normal distribution tables.

Corollary 1 *With the same assumptions of Theorem 2, the number of plaintext blocks required to have a certain success probability* P_S *in an a-bit advantage linear attack is equal to* $c_N|p-1/2|^{-2}$, *where*

$$c_N = \left(\frac{\Phi^{-1}(P_S) + \Phi^{-1}(1-2^{-a-1})}{2}\right)^2 . \tag{14}$$

2.4 Accuracy of the Approximations

In a typical linear attack, N is at least in the order of 2^{30}–2^{40} and p is very close to $1/2$. Hence, the normal distribution can be expected to give an extremely good approximation for the binomial T_i counters and for $X_i = (T_i/N - 1/2)$. As for the normal approximation of the order statistics, it is usually accepted to be a good approximation when n is in the order of hundreds or larger [3]. In our case, $n = 2^m - 1$, hence, we conjecture that the normal distribution will be a good approximation, in particular when $m \geq 16$, as in most linear attacks.

Although it is difficult in general to verify the goodness of the normal approximation for the order statistics, it can be done quite efficiently for the special case $a = m$ (i.e., when the right key is to be ranked the highest). For this case, the probability of success is

$$\begin{aligned} P_S(m) &= \int_0^\infty \left(\int_{-x}^{x} f_W(y)\,dy\right)^{2^m-1} f_0(x)\,dx \\ &= \int_{-2\sqrt{N}|p-1/2|}^{\infty} \left(\int_{-x-2\sqrt{N}|p-1/2|}^{x+2\sqrt{N}|p-1/2|} \phi(y)dy\right)^{2^m-1} \phi(x)dx\,. \end{aligned} \tag{15}$$

We calculated (15) for $m \leq 32$. The results match the results in Table 2 with an error rate of 5% or less. The relatively high error rates occur for $0.1 < P_S < 0.5$. Where $P_S > 0.9$ is of concern, the error rates are less than 1%.

2.5 Discussion on the Results

In this section, we gave three alternative expressions of the success probability in a linear attack, (9), (13), and (15); all assuming that the T_i counters are independent and can be approximated by a normal distribution, and that the linear approximation has a zero bias for wrong keys. (15) is the most accurate among the three, but it is also the most costly to calculate and is limited to $a = m$. (9) is a more general expression, not limited to $a = m$, obtained by the normal approximation to the order statistics. (13) is a simplification of (9), by the observation of $\sigma_q^2 \ll \sigma_0^2$. It gives an expression of the success probability as a function of the advantage a, independent of m, and also gives a formula for calculating the amount of plaintext required for a certain success probability.

We would like to note it again that the probability calculations in this section assume that the linear approximation's bias is zero for all wrong keys, which is the ideal case for the attacker but may not be true in practice. Therefore, the probability calculations here must be taken as an upper bound. We will discuss the issue of wrong key biases in more detail in Section 5.

Finally, we would like to note that the one bit of key information derived in a linear attack from the xor of the key bits on the right-hand side of the approximation is not included in our notation of the advantage a. Counting that bit of information, the advantage of the attack would be $a+1$ bits, if the xored bits are not all included among the derived key bits.

3 Success Probability in Differential Cryptanalysis

In a differential attack, the attacker first finds 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. After having 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.[2] 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, even 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, $k_i, 1 \leq i \leq 2^m - 1$, 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. $W_i, 1 \leq i \leq 2^m - 1$, 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 inital difference. S_N denotes the signal-to-noise ratio, p/p_r.

[2] If a pair suggest no keys, it is certainly a "wrong pair" and can be discarded.

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. In a typical differential attack, 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,

$$\begin{array}{lll} p_0 = p + (1-p)p_r \approx p + p_r, & \mu_0 = p_0 N, & \sigma_0^2 = p_0(1-p_0)N \approx p_0 N, \\ p_W = p_r, & \mu_W = p_W N, & \sigma_W^2 = p_W(1-p_W)N \approx p_W N. \end{array}$$

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

$$P_S = \int_{-\frac{\mu_0 - \mu_q}{\sqrt{\sigma_0^2 + \sigma_q^2}}}^{\infty} \phi(x)\, dx\,, \tag{16}$$

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

$$P_S = \int_{-\frac{\mu_0 - \mu_q}{\sigma_0}}^{\infty} \phi(x)\, dx\,. \tag{17}$$

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

$$\begin{aligned} \frac{-\mu_0 + \mu_q}{\sigma_0} &= \frac{-p_0 N + p_W N + \sqrt{p_W N}\Phi^{-1}(1-2^{-a})}{\sqrt{p_0 N}} \\ &= \frac{-pN + \sqrt{p_r N}\Phi^{-1}(1-2^{-a})}{\sqrt{(p+p_r)N}} \\ &= -\sqrt{pN}\sqrt{\frac{p}{p+p_r}} + \sqrt{\frac{p_r}{p+p_r}}\Phi^{-1}(1-2^{-a}) \\ &= -\sqrt{\mu}\sqrt{\frac{S_N}{S_N+1}} + \sqrt{\frac{1}{S_N+1}}\Phi^{-1}(1-2^{-a})\,. \end{aligned} \tag{18}$$

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,*

$$P_S \; = \; \int_{-\frac{\sqrt{\mu S_N} - \Phi^{-1}(1-2^{-a})}{\sqrt{S_N+1}}}^{\infty} \phi(x)\, dx \,, \tag{19}$$

where $\mu = pN$.

Corollary 2 *With the same assumptions of Theorem 3, the number of plaintext-ciphertext pairs required to have a certain success probability P_S in an a-bit advantage differential attack is*

$$N \; = \; \frac{\left(\sqrt{S_N+1}\Phi^{-1}(P_S) + \Phi^{-1}(1-2^{-a})\right)^2}{S_N}\, p^{-1} \,. \tag{20}$$

A numerical calculation of (19) for $S_N = 1$ and $S_N = 1000$ is given in Table 3 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 chance only for $a < 20$. To have a good success chance for larger values of a as well, 80 or more right pairs would be needed.

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) $S_N = 1$

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) $S_N = 1000$

Table 3: Probability of achieving an a-bit advantage for various values of the expected number of right pairs μ, according to equation (19).

3.3 Accuracy of the Approximations

The normal approximation for the binomial T_0 can be expected to be quite good in general, since typically $p_0(1-p_0)N$ will be at least 4 or higher. However, the same cannot be said for other T_is if S_N is large, which implies $p_W N = \mu/S_N$ will be very small. In those cases, instead of using $\sigma_W \Phi^{-1}(1-2^{-a})$ for μ_q, the actual $\mu_q = F_W^{-1}(1-2^{-a})$ can be used where F_W is the binomial distribution $\mathcal{B}(N, p_W)$. However, this method should be preferred only if a high precision is required, since a numeric calculation of F_W^{-1} would be very costly. Otherwise, if a high precision is not required, we believe the results obtained by the normal approximation are reasonably good, especially considering the fact that the value of μ is dominated mostly by $\Phi^{-1}(P_S)$ rather than $F_W^{-1}(1-2^{-a})$ when S_N is large. When S_N is small, the normal approximation should be good for all T_is, since in that case $\mu = pN$ will be taken higher and $p_W(1-p_W)N$ will be sufficiently large as well.

Regarding the normal approximation for the order statistics, it is usually accepted to give a good approximation for fairly large n, as we discussed in Section 2.4. We have $n = 2^m - 1$; so, we do not expect this approximation to cause any serious problems, especially as long as $m \geq 16$. The goodness of the approximation can be tested efficiently for $a = m$. For this case, a quick analysis, again assuming the independence of the counters and the normal approximation for the binomial distribution, gives,

$$\begin{aligned} P_S(m) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{x} f_W(y)\, dy \right)^{2^m-1} f_0(x)\, dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{x\sqrt{sn+1}+\sqrt{\mu\, sn}} \phi(y) dy \right)^{2^m-1} \phi(x) dx \, . \end{aligned} \tag{21}$$

We calculated (21) for $m \leq 32$. The results match the results in Table 3 with an error rate of less than 4%. As in linear cryptanalysis, the relatively high error rates occur for the smaller values of P_S. For $P_S > 0.90$, the error rate is much less than 1%.

3.4 Discussion on the Results

We gave three expressions of the success probability in differential cryptanalysis, similar to those in linear cryptanalysis. Among them, (21) is the most accurate but is also the most expensive to calculate, and it is limited to $a = m$. (16) is a more general expression, applicable to arbitrary a, m, and assumes the normal approximation for the order statistics. (19) is a simplification of (16) for $\sigma_q^2 \ll \sigma_0^2$, which gives an expression for the success probability independent of m and a formula for calculating the required amount of data for a certain success probability.

4 A Limitation of Ranking Attacks

Although may be an effective alternative on ciphers with DES-like key schedules, attacks that get the right key among the first few thousand candidates rather than as the first one have a very limited applicability on ciphers with a one-way key schedule, where every round key would have to be discovered separately: To attack an inner round key, first all the the outer keys that cover it must be discovered; and they must be discovered fully, not in part. Moreover, a wrong guess for an outer round key will not be detected until all candidates for the inner rounds are tried, at least up to a certain treshold. For example, if 10 is the treshold for the maximum number of keys to be tried at each round of a 16-round cipher, a wrong guess for the last round key will not be known until 10^{15} candidates are tried for the inner rounds. An attack requiring the testing of a few thousand keys at every round to find the right key would be totally infeasible, even more expensive than the exhaustive search.

In the next section, we give an example where a linear attack cannot find the correct round key among the first few hundred candidates, even after analyzing all possible plaintext blocks. The main problem there is the non-negligible bias of the wrong keys, combined with the significant key dependence of the approximation. If such a cipher whose approximations are strongly key dependent also has a one-way key schedule, we believe its linear cryptanalysis will be a mostly theoretical question.

5 Bias of the Right and Wrong Keys

The bias of a linear approximation with a wrong key may not be negligible compared to the bias with the right key, contrary to what is commonly assumed. We discuss two factors that contribute to this phenomenon.

1. **Non-zero bias of wrong keys** A typical linear approximation of a 1-R linear attack [7] is of the form

$$P[i_1, \ldots, i_a] \oplus C[j_1, \ldots, j_b] \oplus F(C, K)[t_1, \ldots, t_c] = K[u_1, \ldots, u_d], \tag{22}$$

   where P, C, and K are the plaintext, ciphertext, and key, and $[i, \ldots, j]$ denotes the xor of some specified bits. F is a function related to the last round function of the cipher and is used in conjunction with $C[j_1, \ldots, j_b]$ to express the $(r-1)$th round output.

   Unlike commonly thought, the approximation with a wrong key k_i substituted for K in F does not result in something random: It will be just another approximation, for an $(r+1)$-round cipher with $C \oplus F(C, k_i)$ as the last round function. The bias of this new approximation, although lower than the original $(r-1)$-round approximation, will be far from zero.

2. **Key dependence of the bias.** If an approximation is significantly key dependent, its bias can be many times higher or lower than its average value, depending on the encryption key in place. Hence, some of the wrong-key biases can plausibly become very close to the right-key bias and may even exceed it due to these fluctuations. We give such an example in Section 5.1.

The effect of the first factor above is not so serious if there is no significant key dependence: Assume that the bias of an approximation becomes μ_W $(\neq 0)$ when a wrong key is substituted.

Then the success probability (12) becomes

$$\int_{-2\sqrt{N}(\mu_0-\mu_W)+\Phi^{-1}(1-2^{-a-1})}^{\infty} \phi(x)\, dx \,, \tag{23}$$

where $\mu_0 = |p - 1/2|$. As discussed above, the effect of a wrong key in the approximation is roughly equivalent to having two extra rounds in the approximation. The decline in the bias in two rounds would normally be sufficient to make μ_W/μ_0 negligibly small. On the other hand, if there is a significant key dependence, the expected bias can be much greater than the average μ_W for a subset of wrong keys, as in the example below.

5.1 An Example

The RC5 encryption algorithm is defined as

$L_1 = L_0 + K_0$
$R_1 = R_0 + K_1$
for $i = 2$ **to** $2r + 1$ **do**
  $L_i = R_{i-1}$
  $R_i = ((L_{i-1} \oplus R_{i-1}) \lll R_{i-1}) + K_i$

where L_i and R_i denote the left and right halves of the text after the ith round, and K_i is the ith round key. All L_i, R_i, K_i are w bits long, which is a variable parameter known as the *word* size. r denotes the number of rounds, which is also a variable parameter, and each iteration of the for loop is known as a half-round. (L_0, R_0) is the plaintext block and (L_{2r+1}, R_{2r+1}) gives the ciphertext.

The best linear approximation of RC5 presently known is

$$R_0[0] \oplus L_{2r}[0] = K_1[0] \oplus K_3[0] \oplus \cdots \oplus K_{2r-1}[0], \tag{24}$$

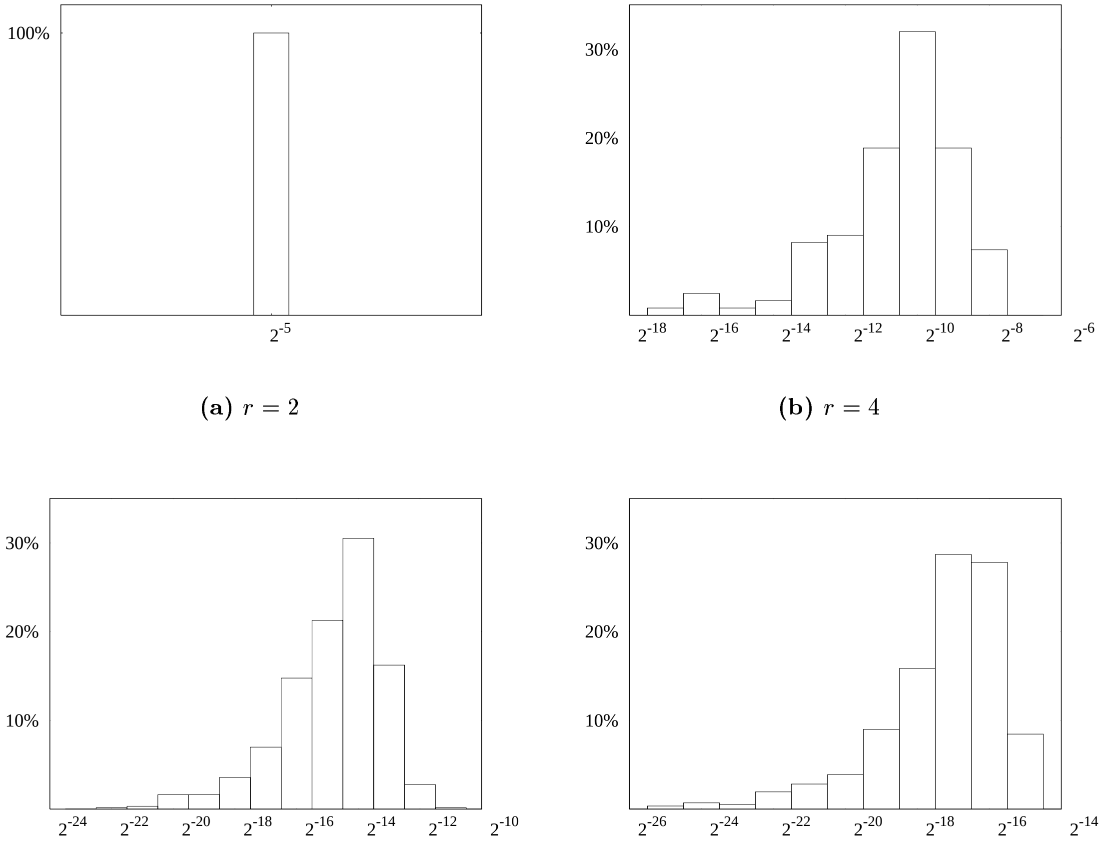
discovered in [5]. Our tests with this approximation have shown that its bias changes significantly depending on the encryption key. Figure 1 summarizes these findings for $w = 16$.

$L_{2r}[0]$ of approximation (24) can be represented in terms of the ciphertext and the last round key as $(R_{2r+1} - K_{2r+1})[\rho] \oplus L_{2r+1}[0]$ where ρ denotes $L_{2r+1} \bmod w$. This substitution yields

$$R_0[0] \oplus (R_{2r+1} - K_{2r+1})[\rho] \oplus L_{2r+1}[0] = K_1[0] \oplus \cdots \oplus K_{2r-1}[0]. \tag{25}$$

Approximation (25) can be used to discover K_{2r+1} by trying all possible key values with a large sample of known plaintexts and then choosing the key which maximizes the bias [10]. However, our experiments show that most of the time there will be a significant number of wrong keys with a bias higher than that of the right key. Figure 2 summarizes the results of these experiments where the bias for each key was computed exhaustively over all 2^{32} plaintext blocks.

The randomness in the ranking for $r = 8$ in Figure 2 is hardly surprising, since the bias $2^{17.2}$ is about the bias of a random approximation, that is, $2^{-\frac{n}{2}-1}$ for an n-bit block



(a) $r = 2$ **(b)** $r = 4$

(c) $r = 6$ **(d)** $r = 8$

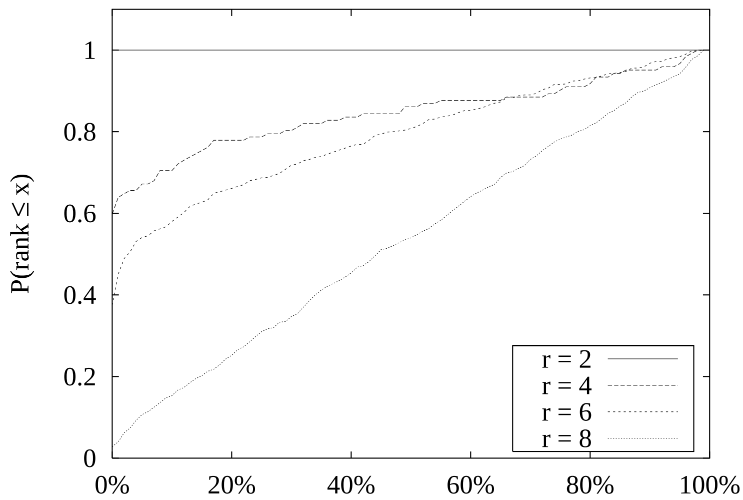
Figure 1: Distribution of the bias of approximation (24) for r rounds. The bias is calculated exhaustively for each key by going over all 2^{32} plaintext blocks. Each plot is based on a sample of more than 500 keys.

cipher [2, 11]. More informative are the cases $2 \leq r \leq 6$, where both an increase in the key dependence and a decline in the ranking of the right key occur concurrently with increasing number of rounds.

For the moment, we leave the formulation of the relation between the key dependence and the ranking of the right key as an open problem.

6 Conclusions

We presented an analytical calculation of the success probability and the data requirement of linear and differential attacks. The derived formulae can be computed very efficiently



(a) Cumulative distribution of the right key rank.

r	2	4	6	8
Bias	2^{-5}	$2^{-10.4}$	$2^{-14.6}$	$2^{-17.2}$
$E[rank]$	0%	16%	22%	48%

(b) The average bias and rank of the right key.

Figure 2: The ranking of the right key among all possibilities for K_{2r+1}, according to their bias in approximation (25), taken over the key samples used for Figure 1. The results show a decline in the ranking of the right key with increasing number of rounds. The ranking becomes completely random for $r = 8$.

and they provide a practical tool for the success probability estimation. We conjecture the approximations and assumptions taken during the analysis to be reasonably good, especially in the case of differential cryptanalysis. The assumption of negligible bias for all wrong keys in linear cryptanalysis is likely to be unrealistic in certain attacks where the approximation's probability is significantly key dependent. The success probability obtained by this assumption can be used as an upper bound, nevertheless. We leave the analysis of the exact relationship between the key dependence of a linear approximation and the ranking of the right key obtained according to that approximation as an open problem.

Acknowledgments

I would like to thank to Ali Bıçak for his help with the experiments and for many helpful comments, and to Burgess Davis for insightful discussions on order statistics.

References

[1] E. Biham and A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard.* Springer-Verlag, 1993.

[2] Florent Chabaud and Serge Vaudenay. Links between differential and linear cryptanalysis. In A. De Santis, editor, *Advances in Cryptology—Eurocrypt'94*, pages 356–365. Springer-Verlag, 1994.

[3] Burgess Davis. Personal communication.

[4] Pascal Junod. On the compexity of Matsui's attack. In *Workshop on Selected Areas in Cryptography (SAC 2001).* Springer-Verlag, 2001.

[5] Burton S. Kaliski Jr. and Yiqun Lisa Yin. On differential and linear cryptanalysis of the RC5 encryption algorithm. In D. Coppersmith, editor, *Advances in Cryptology—Crypto'95*, pages 171–184. Springer-Verlag, 1995.

[6] F. C. Leone, N. L. Nelson, and R. B. Nottingham. The folded normal distribution. *Technometrics*, 3:543–550, 1961.

[7] Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In T. Helleseth, editor, *Advances in Cryptology—Eurocrypt'93*, pages 386–397. Springer-Verlag, 1993.

[8] Mitsuru Matsui. The first experimental cryptanalysis of the Data Encryption Standard. In Y. G. Desmedt, editor, *Advances in Cryptology—Crypto'94*, pages 1–11. Springer-Verlag, 1994.

[9] A. Rényi. *Probability Theory.* American Elsevier Publishing Company, Inc., 1970.

[10] Ali Aydın Selçuk. New results in linear cryptanalysis of RC5. In S. Vaudenay, editor, *Fast Software Encryption, 5th International Workshop*, pages 1–16. Springer-Verlag, 1998.

[11] Ali Aydın Selçuk. On bias estimation in linear cryptanalysis. In *Indocrypt 2000*, pages 52–66. Springer-Verlag, 2000.

A 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

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

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* [6], which we denote by $\mathcal{FN}(\mu, \sigma^2)$. The mean and variance of Y are,

$$\begin{aligned} E(Y) &= \mu(1 - 2\Phi(-\mu/\sigma)) + 2\sigma\phi(\mu/\sigma) \\ Var(Y) &= \mu^2 + \sigma^2 - E(Y)^2. \end{aligned}$$