Analysis of RC6

January 12, 2001

Executive Summary

This report presents the results of a limited evaluation of the block cipher RC6. The evaluation consisted of theoretical derivations and practical experimentations.

No flaws nor weaknesses have been identified in the design which could lead to cryptanalytic attacks with respect to the state-of-the-art. The prediction by the designers that only up to 16 of the 20 rounds can be attacked still stands today. Thus, the security margin for RC6 with the proposed number of rounds remains the same.

The design of RC6 is very close to the design of RC5 which has been subject to public scrutiny for more than five years. We believe that RC6 is a stronger design than that of RC5.

Finally we would like to mention that a longer, concentrated analysis of RC6 might reveal properties which we were not able to detect in this limited time analysis.

Contents

1 Structural features and characteristics

RC6 has a simple structure and description relative to other proposed block ciphers. In the following we refer to [1, 2, 3, 4] for further descriptions and notation. RC6 was one of five finalists for the Advanced Encryption Standard [35]. It consists of two Feistel networks whose data are mixed via data-dependent rotations. The operations in one round of RC6 are the following: two applications of the squaring function $f(x) = x(2x + 1) \bmod 2^{32}$, two fixed 32-bit rotations, two data-dependent 32-bit rotations, two exclusive-ors and two additions modulo 2^{32}. The cipher runs in 20 rounds. RC6 is an evolutionary extension of the block cipher RC5, which received much attention because of its design which is even simpler than that of RC6. Where RC5 works on two 32-bit words, RC6 is extended to operations on four 32-bit words. The relative simple structure of RC5 has allowed for some easy analysis and yet it seems that 16 rounds of RC5 still resists all known attacks well.

The design of RC6 is more complex than that of RC5, and consequently an analysis of the cipher gets more involved. The security of RC6 relies on the strength of data-dependent rotations, the mixed use of exclusive-or operations and modular additions, and on the squaring function f together with the fixed rotations. By removing one or several of these diffusion properties, the resulting cipher is weaker with respect to known attacks. However, it speaks in favor of RC6 that some of these attacks are only theoretical with no practical attacks.

2 Evaluation of security level in terms of differential and linear cryptanalysis

In section we consider differential and linear cryptanalysis of RC6.

Let us first consider the individual components of RC6 and differential attacks. We shall define the difference between two texts as the exclusive-or of the texts and consider differences in 32-bit words. First consider the addition of a subkey modulo 2^{32}. Integer addition of a constant word S to words A and B which only differ in few bits does not necessarily lead to an increase of bit differences in the sums $A + S$ and $B + S$. This may be illustrated by the following special cases: Suppose the words A and B only differ in the most significant bit. Then it follows that $A + S$ and $B + S$ also differ in only the most significant bit. Suppose next that the words A and B only differ in the i-th bit, $i < 31$. Then it can be shown that with probability $\frac{1}{2}$, $A+S$ and $B+S$ also differ in only the i-th bit. If we use the binary representation of words, i.e., $A = a_{w-1}2^{w-1} + \cdots + a_1 2 + a_0$, and similarly for B and S, the binary representation of the sum $Z = A + S$ may be obtained by the formulae

$$z_j = a_j + s_j + \sigma_{j-1} \text{ and } \sigma_j = a_j s_j + a_j \sigma_{j-1} + s_j \sigma_{j-1}, \tag{1}$$

where σ_{j-1} denotes the carry bit and $\sigma_{-1} = 0$ (cf. [37]). Using these formulae one sees that $A + S$ and $B + S$ with probability $\frac{1}{4}$ differ in exactly two (consecutive) bits. Suppose now the words A and B already differ in exactly two

consecutive bits. Then again using the formulae (1) one can see that with probability $\frac{1}{4}$, $A + S$ and $B + S$ differ in exactly one bit and that with probability $\frac{3}{8}$, $A + S$ and $B + S$ differ in exactly two (not necessarily consecutive) bits. Thus with probability $\frac{5}{8}$ the words $A + S$ and $B + S$ differ again in at most two bits if A and B differ in two consecutive bits. Using the formulae (1) one could discuss relations between integer addition and bit differences in a more general setting. However the above suggests that addition of sub keys can only moderately contribute to an avalanche effect of bit differences. The following concerns a relationship between rotations and bit differences in RC6. It is clear that if A and B differ in s bits, then if both words are rotated by the same amount, the resulting texts also differ in exactly s bits. If the texts are rotated by a different amount then the resulting difference becomes harder to predict. These reasonings motivate to consider exclusive-or differences of low Hamming weights, e.g., one or two.

Next let us consider the squaring function $f(x) = x(2x + 1) \bmod 2^{32}$. We use the following notation $\Delta x \xrightarrow{G} \Delta x'$ if texts of differences Δx can result texts of differences $\Delta x'$ after one application of a function G. It follows by easy calculations that there is a non-trivial differential through f of probability one. Assume that x_0 and x_1 are two 32-bit texts different in only the most significant bit. Then $f(x_0)$ and $f(x_1)$ are also different only in the most significant bit. In other words the following differential in hex notation holds with probability one:

$$80000000_x \quad \xrightarrow{f} \quad 8000000_x.$$

Assume next that x_0 and x_1 are two 32-bit texts different in only the second-most significant bit. Then $f(x_0)$ and $f(x_1)$ are also different only in the second-most significant bit with probability $1/2$. In the other cases $f(x_0)$ and $f(x_1)$ are different in both the most significant and in the second-most significant bits. In a similar manner, consider two 32-bit texts different in only the third-most significant bit. Then after the application of the function f, the texts will differ in at most three bits, the three most significant bits. All these differentials can be expressed as a (so-called) truncated differentials which holds with probability one: it holds that if two texts are equal in the s least significant bits, then after the application of the function f, the texts are equal in at least the s least significant bits.

It follows that for all individual components of RC6 there exist differentials of high probabilities. The question is if it is possible to exploit these high probability differentials in an attack. For this to happen, it seems one must consider texts of differences with a low Hamming weight, but such that all data-dependent rotations are equal for the two texts. The authors themselves have analysed RC6 extensively with respect to such differentials. In [3] results are given which show that there are only a limited number of possible differentials satisfying these constraints. These differentials however only allow for attacks on RC6 reduced to 12 rounds or less. Since it is recommended to use 20 rounds, it is believed that with respect to these differential attacks the security margin

is more than sufficient. Let us next consider differentials where the difference is defined by subtraction modulo 2^{32}. Thus, if x_0 and x_1 are two 32-bit texts, the difference is defined $\Delta x = x_0 - x_1 \bmod 2^{32}$. It follows that such differentials have probability one through a modular addition of a subkey modulo 2^{32}. Let us next consider the squaring function f. Let x_0 and x_1 be two 32-bit texts of difference α. Then the difference in the texts after the application of the function f is

$$\begin{aligned} f(x_0) - f(x_1) &= x_0(2x_0+1) - x_1(2x_1+1) \\ &= (x_1+\alpha)(2(x_1+\alpha)+1) - x_1(2x_1+1) \\ &= 4\alpha x_1 + 2\alpha^2 + \alpha, \end{aligned}$$

thus the difference depends on both α and x_1. However, since x_1 appears in first degree, a second-order differential will not depend on neither x_0 nor on x_1. To see this, consider x_0 and x_1 as above and two texts x_2 and x_3 of difference α, such that the difference between x_1 and x_3 is some value β. The four texts x_0, x_1, x_2, and x_3 form a second-order differential. The value of the differential through the function f is computed as follows.

$$\begin{aligned} f(x_0) - f(x_1) - (f(x_2) - f(x_3)) &= 4\alpha x_1 + 2\alpha^2 + \alpha - (4\alpha x_3 + 2\alpha^2 + \alpha), \\ &= 4\alpha(x_1 - x_3) \\ &= 4\alpha\beta, \end{aligned}$$

thus not depending on the input texts, only on their differences. This means that through the function f there are second-order differentials of probability one. The problem in using these second-order differentials is first the data-dependent rotations. If one should be able to iterate such differentials through several rounds of RC6, then it seems one has to assume that the rotations for all four texts always are the same. Thus, the probability of such differentials are expected to be applicable only for a few number of rounds after which the probabilities will be very low.

Let us next consider linear cryptanalysis and examine the individual components of RC6. First of all, it follows that the addition of a subkey modulo 2^{32} introduces, in general, carry-bits which complicate linear approximations. The above considerations about the mixed use of exclusive-ors and modular additions illustrates this fact. However, there are linear relations through a modular addition of probability one. As before, let A be 32 bits of data and let S be a 32-bit subkey. Then it follows that the least significant bits of A and $A+S$, have maximum bias, since the least significant bit of S is a constant. Consider next the squaring function f. It follows that the least significant bits of x and $f(x)$ are equal. Thus, there are linear approximations through f of probability one. This situation is quite similar to the one for differential cryptanalysis, where the most significant bits play the important role. The problem of iterating linear approximations through several rounds is similarly difficult as for differential cryptanalysis. The linear cryptanalysis performed by the authors conjecture that RC6 is vulnerable to this attack for a maximum of 16 rounds, in which

case the attack will be very involved and require a huge and unrealistic amount of plaintexts and their corresponding ciphertexts. In the next section we shall consider a correlation attack which is similar to the best linear attack discovered by the authors.

3 Evaluation of security level in terms of other cryptanalysis

In this section we consider other methods of cryptanalysis of RC6.

First of all, there are trivial attacks which apply to all block ciphers. An exhaustive key search will take 2^k operations to succeed, where k is the key size. Also, the "matching ciphertext attack" applies in ECB and CBC mode, but requires about $2^{n/2}$ ciphertext blocks to succeed with good probability, where n is the block size. With $n = 128$ as in RC6, 2^{64} ciphertext blocks are required after which an attacker would be able to deduce information about the plaintext blocks.

Higher order differentials. This attack applies to ciphers which uses non-linear components of a low algebraic degree. The non-linear data-dependent rotations complicate the higher order differential attacks. It is believed that s-order differential attacks are less serious for RC6 for increasing values of s. We discussed higher order differentials already earlier in this report.

The slide attacks, the integral attacks, the non-surjective attacks and the "mod n" attacks do not seem applicable to RC6 . The latter might be applicable to modified variants of RC6, which will be discussed later.

The interpolation attacks apply to ciphers which use simple mathematical functions only. RC6 uses mathematical functions in the squaring function, however the mixed use of exclusive-ors and modular addition together with data-dependent rotations seems to have a good effect in thwarting the interpolation attacks.

The key-schedule of RC6 does not seem to allow for related-key attacks. Since the round keys are encrypted in relatively many rounds using an encryption routine similar to that of the encryption algorithm itself, it is unlikely that any easily identified weak keys or pairs of related keys exist.

In [26, 17] correlation attacks on RC6 were presented. In the following we shall outline these attacks. The focus will be on the analysis of [26] since this author is a co-author of this work.

3.1 Correlation attack

In the approach of [26] one fixes each of the least significant five bits in the first word A and the third word C of the plaintexts and investigates the statistics of the 10-bit integer obtained by concatenating each of the least significant five bits in the first and third words, A'' and C'', two rounds later. This is motivated by the fact that the least significant five bits in A and C altogether are not changed by the xor and data dependent rotation if both rotation amounts are

zero. More generally, one can expect a bias for amounts smaller than five. This leads to a strong bias which can be iterated over many rounds, just as in linear approximations. Once can consider small rotation amounts as well as the zero rotation and also rotations near zero from the negative, like 30 or 31, prove to be useful as well.

Next we summarise the nonrandomness of r-round versions of RC6. The analysis is based on systematic experiments on increasing numbers of rounds of RC6 with varying word length w. The method is used to demonstrate that detecting and quantifying nonrandomness is experimentally feasible up to 6 rounds of RC6.

The tool for the tests is the χ^2 statistic of the integer of size ten bits as obtained by concatenating the least significant five bits in words A'' and C'' every two rounds later.

It has been shown that these tests make it possible to distinguish RC6 with a certain number of rounds from a permutation randomly chosen from the set of all permutations. Table 1 lists the result of tests implemented for RC6 with 128-bit blocks with 3 and 5 rounds. It follows that $2^{13.8}$ texts are sufficient to distinguish the 3-round encryption permutation from a randomly chosen permutation in 90% of the cases. It was estimated that for RC6 with $3+2r$ rounds similar results will hold using $2^{13.8+r\times 16.2}$ texts, which was confirmed by tests implemented on RC6 with 5 rounds.

Also, it was estimated that for keys where the least significant five bits of each of the two subkeys in every second round are zeros, the attack improves with more than a factor of two for each 2 rounds. This leads to the estimate that for one in 2^{80} keys, 17 rounds of RC6 with 128-bit blocks can be distinguished from a randomly chosen permutation.

r	#Texts	Comments
3	2^{13}	
3	$2^{13.8}$	
3	2^{14}	
5	2^{29}	
5	2^{30}	
7	$2^{46.2}$	Estimated.
9	$2^{62.4}$	Estimated.
11	$2^{78.6}$	Estimated.
13	$2^{94.8}$	Estimated.
15	$2^{111.0}$	Estimated.
17	$\leq 2^{118}$	For 1 in every 2^{80} keys.

Table 1: Complexities for distinguishing RC6 with 128-bit blocks and r rounds from a random function.

Also, it was shown that these findings can be used in key-recovery attacks on RC6. Table 2 lists the estimated complexities of the key-recovery attack for RC6 with up to 12 rounds for all keys, up to 14 rounds for 192-bit key variants,

r	#Texts	Work	Memory	Comment
12	2^{94}	2^{119}	2^{42}	
14	2^{110}	2^{135}	2^{42}	
14	2^{108}	2^{160}	2^{74}	
15	2^{119}	2^{215}	2^{138}	
16	2^{118}	2^{171}	2^{74}	1 in 2^{60} keys

Table 2: Complexities for key-recovery attacks on RC6 with 128-bit blocks and r rounds. One unit in "Work" is the time to increment one counter.

up to 15 rounds for 256-bit key variants, and up to 16 rounds for some weak keys.

The analysis of [17] is comparable to above outlined correlation attack.

3.1.1 New analysis

In this section we report on a new analysis conducted on RC6 for this report. The idea is that instead of estimating the nonuniformness of five bits from each of the two Feistel halves in RC6, that is, in total ten bits, there is the possibility that it could be advantageous to consider instead less or more bits. To analyse this approach we implemented a series of tests on RC6. Let us call the above non-uniformness a 10-bit correlation. Then we shall consider also 8-bit correlations and 12-bit correlations.

First a number of tests were run on RC6 with 128-bit blocks as is the proposed version. By considering the uniformness of 10 bits of the ciphertexts after 2 rounds, one needs to generate about 2^{12} texts to detect nonrandomness. However, in a 12-bit correlation there are 2^{12} possible values of the bits considered, therefore a χ^2-tests is likely to be unreliable (it is usually recommended to use 5 times as many samples as there are possible values). Moreover, tests on a 2-round version might not accuratly measure the effect multiple round correlations might have. In a 4 round version of RC6 one would need to generate 2^{28} texts to see the effect. As we planned to analyse both the 8, 10, and 12-bit correlations and since in each case one would need to do several tests, this test was regarded too time-consuming.

Instead we implemented RC6 working on 16-bit words instead of 32-bit words. One advantage of RC6 is that it scales easily up and down in word sizes. In an RC6 version with 64-bit blocks the correlation attack as outlined above would measure the nonuniformness in 8 bits. Without going into too many details, the expected value of the χ^2 statistic is in this case 255. The number of texts needed to measure the effect of nonrandomness is much lower for this version of RC6.

Consider Table 3, which lists the results of χ^2 tests on 4 rounds of RC6 with 64-bit blocks. As can be seen, there does not seem to be a big effect in considering less than or more than eight bits. The results in the table show that the correlations are comparable. Therefore, it seems we can conclude from

#Texts	6-bit	8-bit	10-bit
2^{20}	66	253	1014
2^{22}	68	266	1053

Table 3: The χ^2-values for RC6 with 64-bit blocks and 4 rounds. Expected χ^2 for random 6-bit, 8-bit, and 10-bit functions are 63, 255, and 1023, respectively.

this analysis that 8-bit and 12-bit correlations would not prove a substantial improvement as compared to the known 10-bit correlations.

3.2 Modified variants of RC6

In [3] the designers of RC6 considered the security of some weakened variants of the algorithm. The motivation for this is to illustrate that the components left out are significant for the security of the algorithm. In the following we shall extend the number of possible modifications of RC6.

1. Remove the fixed rotations.
2. Replace the squaring function f by the identity function.
3. Change the exclusive-ors to modular additions.
4. Change the modular additions to exclusive-ors.
5. Remove the data-dependent rotations.

In [3] the first two options were considered. It was shown that RC6 wihtout the fixed rotations result in a weaker cipher, the main reason being that the data-dependent rotation depend on only five bits of the inputs to f. Also, it was shown that removing the squaring function would give a much weaker cipher. This variant is somewhat comparable to two parallel runs of the RC5 algorithm.

Let us consider the third modification, that is changing the exclusive-ors to modular additions. Then since the squaring function is "just" a number of modular additions modulo 2^{32}, the strength of the cipher would lie in the good interation of rotations and modular additions. It has been shown in [20] that such ciphers are susceptible to the so-called "mod-n attacks". Although, it is not at all clear how such an attack would work on this modification of RC6, it is felt that the "mod n" attacks or variants of these might be applicable.

Let us consider the fourth modification, that is, changing the modular additions to exclusive-ors. In this case, it is believed that the correlation attack outlined above has an improved performance. The modular additions of subkeys introduces some carry bits which add to the confusion and decreases the correlations used in the attack.

Let us next consider the fifth modification, that is, removing the data-dependent rotations. In this case, it is strongly believed that there exist efficient differential and/or linear attacks.

The above analysis illustrates that the components are all necessary for RC6 with 20 or less rounds. It is, of course, an impossible task to judge whether the components are sufficient in constructing a secure cipher. So far noone has been able to disprove this conjecture.

4 Survey of previous results on RC6

The known results on RC6 today are first and foremost the first analysis by the designers [3]. In [10] the designers elaborate on the analysis of the simplified variants of RC6, and [11] discusses some differential properties of the data-dependent rotations.

The only known results not from the designers are those of [26, 17, 38]. The first two were already discussed above, the third one concerns scenarios where RC6 is used as a hash function. The author reports on some almost related keys, but to the best of our knowledge these findings are only of theoretical interest, if at all.

A Block Ciphers in General

In the following we give a compressed overview of the state-of-the-art of block cipher cryptanalysis, and outline the following known attacks.

1. Exhaustive Key Search
2. Matching Ciphertext Attacks
3. Differential Cryptanalysis
4. Truncated Differential Attacks
5. Higher-order Differential Attacks
6. Linear Cryptanalysis
7. Related-key Attacks
8. Non-surjective Attacks
9. Interpolation Attacks
10. Mod-n Attacks
11. Slide Attacks
12. Integral Attacks

A.1 Exhaustive key search

This attack needs only a few known plaintext-ciphertext pairs. An attacker simply tries all keys, one by one, and checks whether the given plaintext encrypts to the given ciphertext. For a block cipher with a k-bit key and n-bit blocks the number of pairs of texts needed to determine the key uniquely is approximately $\lceil k/n \rceil$. Also, if the plaintext space is redundant, e.g., consists of English or Japanese text, the attack will work if only some ciphertext blocks is available. The number of ciphertext blocks needed depends on the redundancy of the language.

A.2 The matching ciphertext attack

The *matching ciphertext attack* is based on the fact that for block ciphers of m bits used in the modes of operations for the DES [34] after the encryption of $2^{m/2}$ blocks, equal ciphertext blocks can be expected and information is leaked about the plaintexts [12, 22, 32].

A.3 Differential cryptanalysis

The most well-known and general method of analysing conventional cryptosystems today is *differential cryptanalysis*, published by Biham and Shamir in 1990. Differential cryptanalysis is universal in the sense that it can be used against any cryptographic mapping which is constructed from iterating a fixed round function. One defines a **difference** between two bit strings, X and X' of equal length as

$$\Delta X = X \otimes (X')^{-1}, \tag{2}$$

where $\otimes$ is the group operation on the group of bit strings used to combine the key with the text input in the round function and where $(X)^{-1}$ is the inverse element of X with respect to $\otimes$. The idea behind this is, that the differences between the texts before and after the key is combined are equal, i.e., the difference is independent of the key. To see this, note that

$$(X \otimes K) \otimes (X' \otimes K)^{-1} = X \otimes K \otimes K^{-1} \otimes X'^{-1} = X \otimes (X')^{-1} = \Delta X.$$

In a differential attack one exploits that for certain input differences the distribution of output differences of the non-linear components is non-uniform.

Definition 1 *An s-round* characteristic *is a series of differences defined as an $s+1$-tuple $\{\alpha_0, \alpha_1, \ldots, \alpha_s\}$, where $\Delta P = \alpha_0$, $\Delta C_i = \alpha_i$ for $1 \leq i \leq s$.*

Here ΔP is the difference in the plaintexts and ΔC_i is the difference in the ciphertexts after i rounds of encryption. Thus, the characteristics are lists of expected differences in the intermediate ciphertexts for an encryption of a pair of plaintexts. In essence one specifies a characteristic for a number of rounds and searches for the correct key in the remaining few rounds. In some attacks it is not necessary to predict the values $\alpha_1, \ldots, \alpha_{s-1}$ in a characteristic. The pair (α_0, α_s) is called a *differential*. The complexity of a differential attack is approximately the inverse of the probability of the characteristic or differential used in the attack.

A.4 Truncated differentials

For some ciphers it is possible and advantageous to predict only the values of parts of the differences after each round of the cipher. The notion of truncated differentials was introduced by Knudsen [24]:

Definition 2 *A differential that predicts only parts of an n-bit value is called a* truncated differential. *More formally, let (a, b) be an i-round differential. If a' is a subsequence of a and b' is a subsequence of b, then (a', b') is called an i-round truncated differential.*

A truncated differential can be seen as a collection of differentials. As an example, consider an n-bit block cipher and the truncated differential (a', b), where a' specifies the least $n' < n$ significant bits of the plaintext difference and b specifies the ciphertext difference of length n. This differential is a collection of all $2^{n-n'}$ differentials (a, b), where a is any value, which truncated to the n' least significant bits is a'.

A.5 Impossible differentials

A special type of differentials are those of probability zero. The attack was first applied to the cipher DEAL [25] and later to Skipjack [7]. The main idea is to specify a differential of probability zero over some number of rounds in the attacked cipher. Then by guessing some keys in the rounds not covered by the differential one can discard a wrong value of the key if it would enable the cipher to take on the differences given in the differential.

A.6 Higher-order differentials

An sth-order differential is defined recursively as a (conventional) differential of the function specifying an $(s-1)$st order differential. In order words, an sth order differential consists of a collection of 2^s texts of certain pairwise, predetermined differences. We refer to [28, 24] for a more precise definition of higher order differentials.

In most cases one considers differences induced by the exclusive-or operation and the field of characteristic 2. The *nonlinear order* of a function $f : GF(2^n) \to GF(2^n)$ is defined as follows. Let the output bits y_j be expressed as multivariate polynomials $q_j(x) \in GF(2)[x_1, \ldots, x_n]$, where $x_1, \ldots, x_n$ are the input bits. The nonlinear order of f is then defined to be the minimum total degree of any linear combination of these polynomials. The higher order differential attacks exploit the following result.

Corollary 1 *Let $f : GF(2^n) \to GF(2^n)$ be a function of nonlinear order d. Then any dth order differential is a constant. Consequently, any $(d+1)$st order differential is zero.*

The boomerang attack [39] can be seen as a special type of a second-order differential attack. This variant applies particularly well to ciphers for which one particular (first-order) differential applies well to one half of the cipher, and where another particular (first-order) differential applies well to the other half of the cipher.

A.7 Linear cryptanalysis

Linear cryptanalysis was proposed by Matsui in 1993 [29]. A preliminary version of the attack on FEAL was described in 1992 [31]. Linear cryptanalysis [29] is a known plaintext attack in which the attacker exploits linear approximations of some bits of the plaintext, some bits of the ciphertext and some bits of the secret key. In the attack on the DES (or on DES-like iterated ciphers) the linear approximations are obtained by combining approximations for each round under the assumption of independent round keys. The attacker hopes in this way to find an expression

$$(P \cdot \alpha) \oplus (C \cdot \beta) = (K \cdot \gamma) \tag{3}$$

which holds with probability $p_L \neq \frac{1}{2}$ over all keys [29], such that $|p_L - \frac{1}{2}|$, called the bias, is maximal. In (3) $P, C, \alpha, \beta, \gamma$ are m-bit strings and '$\cdot$' denotes the dot product. The bit strings α, β, γ are called *masks*.

Definition 3 *An s-round* linear characteristic *is a series of masks defined as an $(s+1)$-tuple $\{\alpha_0, \alpha_1, \ldots, \alpha_s\}$, where α_0 is the mask of the plaintexts and α_i is the mask of the ciphertexts after i rounds of encryption for $1 \leq i \leq s$.*

As for differential cryptanalysis one specifies a linear characteristics for a number of rounds and searches for the keys in the remaining rounds, we refer to [29] for more details. A linear attack needs approximately about b^{-2} known plaintexts to succeed, where b is the bias of the linear characteristic used.

Also, the concepts of linear hulls, the analogue to differentials as opposed to characteristics in differentials cryptanalysis, has been defined in [33].

Finally, in [30] it has been shown that if one defines the quantity $q = (2p-1)^2$ where p is the probability of a linear characteristic or hull, then when combining several linear characteristics one can multiply their q values to get the q-value of the combination. Sometimes the q values are referred to as the "linear probability", which is somewhat misleading, but nevertheless seems to be widely used.

A.8 Mod n cryptanalysis

In [20] a generalisation of the linear attacks is considered. This attack is applicable to ciphers for which some words (in some intermediate ciphertext) are biased modulo n, where n typically is a small integer. It has been shown that ciphers which uses only bitwise rotations and additions modulo 2^{32} are vulnerable to these kinds of attacks.

A.9 Related-key attacks

There are several variants of this attack depending on how powerful the attacker is assumed to be.

1. Attacker gets encryptions under one key.
2. Attacker gets encryptions under several keys.
   (a) Known relation between keys.
   (b) Chosen relation between keys.

Knudsen used the methods of 1 by giving a chosen plaintext attack of the first kind on LOKI'91 [21], reducing an exhaustive key search by almost a factor of four. The concept "related-key attack" was introduced by Biham [6], who also introduced the attack scenarios of 2, where the encryptions under several keys are requested. Knudsen later described a related key attack on SAFER K [23] and Kelsey, Schneier, and Wagner [19] applied the related key attacks to a wide range of block ciphers. It may be argued that the attacks with a chosen relation

between the keys are unrealistic. The attacker need to get encryptions under several keys, in some attacks even with chosen plaintexts. However there exist realistic settings, in which an attacker may succeed to obtain such encryptions. Also, there exists quite efficient methods to preclude the related key attacks [19, 16].

A.10 Interpolation attack

In [18] Jakobsen and Knudsen introduced the interpolation attack on block ciphers. The attack is based on the following well-known formula. Let R be a field. Given $2n$ elements $x_1, \ldots, x_n$, $y_1, \ldots, y_n \in R$, where the x_is are distinct. Define

$$f(x) = \sum_{i=1}^{n} y_i \prod_{1 \le j \le n, j \ne i} \frac{x - x_j}{x_i - x_j}. \tag{4}$$

$f(x)$ is the only polynomial over R of degree at most $n-1$ such that $f(x_i) = y_i$ for $i = 1, \ldots, n$. Equation (4) is known as the *Lagrange interpolation formula* (see e.g.,[9, page 185]). In the *interpolation attack* an attacker constructs polynomials using pairs of plaintexts and ciphertexts. This is particularly easy if the components in the cipher can be expressed as easily described mathematical functions. The idea of the attack is, that if the constructed polynomials have a small degree, only few plaintexts and their corresponding ciphertexts are necessary to solve for the (key-dependent) coefficients of the polynomial, e.g., using Lagrange's interpolation. To recover key bits one expresses the ciphertext before the last round as a polynomial of the plaintext.

A.11 Non-surjective attack

In [36] Rijmen-Preneel-De Win described the non-surjective attack on iterated ciphers. It is applicable to Feistel ciphers where the round function is not surjective and therefore statistical attacks become possible. In a Feistel cipher one can compute the exclusive-or of all outputs of the round functions from the plaintexts and the corresponding ciphertexts. Thus, if the round functions are not surjective this gives information about intermediate values in the encryptions, which can be used to get information about the secret keys.

A.12 Slide attacks

In [8] the "slide attacks" were introduced, based on earlier work in [6, 21]. In particular it was shown that iterated ciphers with identical round functions, that is, equal structures plus equal subkeys in the rounds, are susceptible to slide attacks. Let $F_r \circ F_{r-1} \circ \cdots \circ F_1$ denote an r-round iterated cipher, where all F_is are identical. The attacker tries to find pairs of plaintext P, P^* and their corresponding ciphertexts C, C^*, such that $F_1(P) = P^*$ and $F_r(C) = C^*$. Subsequently, an attacker has twice both the inputs and outputs of one round of the cipher. If the round function is simple enough, this can lead to very

efficient attacks. To find such pairs of texts, one can in the worst case apply the birthday paradox, such that one such pair is expected from a collection of $2^{n/2}$ texts, where n is the block size.

A.13 Integral Attacks

These attacks are sometimes referred to as the "Square attack", since it was first applied to the block cipher Square [14, 13]. The attack on Square slightly modified also applies to the block ciphers Crypton and Rijndael [15].

In [27] these attacks are generalised under the name of "integral cryptanalysis". In differential attacks one considers differences of texts, in integral cryptanalysis one considers sums of texts. In ciphers where all nonlinear functions are bijective, it is sometimes possible to predict a sum of texts, even in the cases where differential attacks are not applicable. The main observations are that in a collection of texts which in a particular word take all values exactly equally many times, the value of the words after a bijective function also take all values exactly equally many times. Also, assume that s words have this property and that in the cipher a linear combination of the s words are computed (with respect to the group operation considered). Then it is possible to determine also the sum of all linear combinations in a collection of texts. This attack is still today the best attack reported on Rijndael which has been the selected for the Advanced Encryption Standard.

References

[1] RSA Security. The RC6 Block Cipher. Cryptographic Technique Specifications.

[2] RSA Security. The RC6 Block Cipher. Self Evaluation Report.

[3] S. Contini, R.L. Rivest, M.J.B. Robshaw and Y.L. Yin. The Security of the RC6 Block Cipher. v.1.0, August 20, 1998. Available at `www.rsa.com/rsalabs/aes/`.

[4] R.L. Rivest, M.J.B. Robshaw, R. Sidney and Y.L. Yin. The RC6 Block Cipher. v1.1, August 20, 1998. Available at `www.rsa.com/rsalabs/aes/`.

[5] S. Contini, R.L. Rivest, M.J.B. Robshaw, and Y.L. Yin. Some Comments on the First Round AES Evaluation of RC6. Available at `http://csrc.nist.gov/encryption/aes/round1/pubcmnts.htm`.

[6] E. Biham. New types of cryptanalytic attacks using related keys. In T. Helleseth, editor, *Advances in Cryptology: EUROCRYPT'93, LNCS 765*, pages 398–409. Springer Verlag, 1993.

[7] E. Biham, A. Biryukov, and A. Shamir. Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. In J. Stern, editor, *Advances in*

Cryptology: EUROCRYPT'99, LNCS 1592, pages 12–23. Springer Verlag, 1999.

[8] A. Biryukov and D. Wagner. Slide attacks. In L. R. Knudsen, editor, *Fast Software Encryption, Sixth International Workshop, Rome, Italy, March 1999, LNCS 1636*, pages 245–259. Springer Verlag, 1999.

[9] P.M. Cohn. *Algebra, Volume 1.* John Wiley & Sons, 1982.

[10] S. Contini, R.L. Rivest, M.J.B. Robshaw, and Y.L. Yin. Improved analysis of some simplified variants of RC6. In L. Knudsen, editor, *Fast Software Encryption, Sixth International Workshop, Rome, Italy, March 1999, LNCS 1636*, pages 1–15. Springer Verlag, 1999.

[11] S. Contini and Y.L. Yin. On differential properties of data dependent rotations and their use in Mars and RC6. Presented at the 2nd AES conference, see `www.nist.gov/aes`.

[12] D. Coppersmith, D.B. Johnson, and S.M. Matyas. Triple DES cipher block chaining with output feedback masking. Technical Report RC 20591, IBM, October 1996. Presented at the rump session of CRYPTO'96.

[13] J. Daemen, L. Knudsen, and V. Rijmen. Linear frameworks for block ciphers. *Design, Codes, and Cryptography.* To appear.

[14] J. Daemen, L. Knudsen, and V. Rijmen. The block cipher Square. In E. Biham, editor, *Fast Software Encryption, Fourth International Workshop, Haifa, Israel, January 1997, LNCS 1267*, pages 149–165. Springer Verlag, 1997.

[15] J. Daemen and V .Rijmen. AES proposal: Rijndael. Submitted as an AES Candidate Algorithm. Available from `http://www.nist.gov/aes`.

[16] I.B. Damgård and L.R. Knudsen. Two-key triple encryption. *The Journal of Cryptology*, 11(3):209–218, 1998.

[17] H. Gilbert, H. Handschuh, A. Joux, and S. Vaudenay. A Statistical Attack on RC6. In B. Schneier, editor, *Fast Software Encryption, Seventh International Workshop.* Springer Verlag, 2001. To appear.

[18] T. Jakobsen and L. Knudsen. The interpolation attack on block ciphers. In E. Biham, editor, *Fast Software Encryption, Fourth International Workshop, Haifa, Israel, January 1997, LNCS 1267*, pages 28–40. Springer Verlag, 1997.

[19] J. Kelsey, B. Schneier, and D. Wagner. Key-schedule cryptanalysis of IDEA, G-DES, GOST, SAFER, and triple-DES. In Neal Koblitz, editor, *Advances in Cryptology: CRYPTO'96, LNCS 1109*, pages 237–251. Springer Verlag, 1996.

[20] J. Kelsey, B. Schneier, and D. Wagner. Mod n cryptanalysis, with applications against RC5P and M6. In L. Knudsen, editor, *Fast Software Encryption, Sixth International Workshop, Rome, Italy, March 1999, LNCS 1636*, pages 139–155. Springer Verlag, 1999.

[21] L.R. Knudsen. Cryptanalysis of LOKI'91. In J. Seberry and Y. Zheng, editors, *Advances in Cryptology, AusCrypt 92, LNCS 718*, pages 196–208. Springer Verlag, 1993.

[22] L.R. Knudsen. *Block Ciphers – Analysis, Design and Applications.* PhD thesis, Aarhus University, Denmark, 1994.

[23] L.R. Knudsen. A key-schedule weakness in SAFER K-64. In Don Coppersmith, editor, *Advances in Cryptology - CRYPTO'95, LNCS 963*, pages 274–286. Springer Verlag, 1995.

[24] L.R. Knudsen. Truncated and higher order differentials. In B. Preneel, editor, *Fast Software Encryption - Second International Workshop, Leuven, Belgium, LNCS 1008*, pages 196–211. Springer Verlag, 1995.

[25] L.R. Knudsen. DEAL - a 128-bit block cipher. Technical Report 151, Department of Informatics,University of Bergen, Norway, February 1998. Submitted as an AES candidate by Richard Outerbridge.

[26] L.R. Knudsen and W. Meier. Correlations in RC6 with a reduced number of rounds. In B. Schneier, editor, *Fast Software Encryption, Seventh International Workshop*. Springer Verlag, 2001. To appear.

[27] L.R. Knudsen and D. Wagner. Integral cryptanalysis. In preparation, 2001.

[28] X. Lai. Higher order derivatives and differential cryptanalysis. In R. Blahut, editor, *Communication and Cryptography, Two Sides of One Tapestry.* Kluwer Academic Publishers, 1994. ISBN 0-7923-9469-0.

[29] M. Matsui. Linear cryptanalysis method for DES cipher. In T. Helleseth, editor, *Advances in Cryptology - EUROCRYPT'93, LNCS 765*, pages 386–397. Springer Verlag, 1993.

[30] M. Matsui. New structure of block ciphers with provable security against differential and linear cryptanalysis. In D. Gollman, editor, *Fast Software Encryption, Third International Workshop, Cambridge, UK, February 1996, LNCS 1039*, pages 205–218. Springer Verlag, 1996.

[31] M. Matsui and A. Yamagishi. A new method for known plaintext attack of FEAL cipher. In R. Rueppel, editor, *Advances in Cryptology - EUROCRYPT'92, LNCS 658*, pages 81–91. Springer Verlag, 1992.

[32] U.M. Maurer. New approaches to the design of self-synchronizing stream ciphers. In D.W. Davies, editor, *Advances in Cryptology - EUROCRYPT'91, LNCS 547*, pages 458–471. Springer Verlag, 1991.

[33] K. Nyberg. Linear approximations of block ciphers. In A. De Santis, editor, *Advances in Cryptology - EUROCRYPT'94, LNCS 950*, pages 439–444. Springer Verlag, 1995.

[34] National Bureau of Standards. DES modes of operation. Federal Information Processing Standard (FIPS), Publication 81, National Bureau of Standards, U.S. Department of Commerce, Washington D.C., December 1980.

[35] National Institute of Standards and Technology. Advanced encryption algorithm (AES) development effort. `http://www.nist.gov/aes`.

[36] V. Rijmen, B. Preneel, and E. De Win. On weaknesses of non-surjective round functions. *Designs, Codes, and Cryptography*, 12(3):253–266, 1997.

[37] R.A. Rueppel. *Analysis and Design of Stream Ciphers*. Springer Verlag, 1986.

[38] M-J.O. Saarinen. A Note Regarding the Hash Function Use of MARS and RC6. Available at `www.nist.gov/aes`.

[39] D. Wagner. The boomerang attack. In L. R. Knudsen, editor, *Fast Software Encryption, Sixth International Workshop, Rome, Italy, March 1999, LNCS 1636*, pages 156–170. Springer Verlag, 1999.