Analysis of Hierocrypt-3

January 12, 2001

Executive Summary

This report presents the results of a limited time evaluation of the block cipher Hierocrypt-3.

No flaws nor weaknesses have been identified in the design which could lead to cryptanalytic attacks with respect to the state-of-the-art.

This report contains attacks which are better than the attacks reported by the designers. However, the security margin for Hierocrypt-3 with the proposed number of rounds seems to be adequate for many years still with respect to the known attacks today.

The design of Hierocrypt-3 is very close to the design of Rijndael which has been selected as the Advanced Encryption Standard in the U.S.A. We believe that in case there should be reported new attacks on Rijndael this could also apply to Hierocrypt-3 and the other way around, at least to a certain degree.

Finally we mention that this report is the result of a limited time of review, and the analysis was performed without access to computer code implementing the block cipher. A longer, concentrated analysis might reveal properties of Hierocrypt-3 which we were not able to detect. What speaks in favor of Hierocrypt-3 is its simple design which facilitates for an easy analysis. It is easy relatively to other designs to get convinced about the strength of Hierocrypt-3 against differential and linear cryptanalysis.

Contents

1 Structural features and characteristics

Hierocrypt-3 is an iterated block cipher with 128-bit blocks and allows for three different key sizes to be compliant with the AES [30].

Hierocrypt-3 borrows elements from the block cipher Rijndael [10] but there are differences. Hierocrypt-3 defines a 32-bit S-box XS from an 8-bit S-box S and a linear transformation MDS_L. The 32-bit input to XS is split into four bytes, and each byte is evaluated through S. The 32-bit output of S is input to MDS_L, which is constructed from the parity-check matrix of an MDS-code. The four bytes in the output of MDS_L is then input again to S, and the four output bytes form the output of XS. Before each evaluation of S a key byte is combined to the input via the exclusive-or operation. The linear transformation MDS_H is a permutation on 16 bytes constructed from the parity-check matrix of an MDS-code. Now we can describe the outline of the cipher. The 128-bit plaintext is split into four words of each 32 bits. Each word is input to XS, and then all four words are input to MDS_H. This is repeated five, six or seven times depending on the key length. Finally an output transformation is applied which consists of four applications of the XS S-box together with a final addition of some subkeys. The output transformation allows for similar encryption and decryption routines. We shall call the application of four XS boxes the "XS-layer", the linear transformation the "MDS_H-layer", the application of 16 times the S-box S the "S-box layer" and finally the application of four times the MDS_L-function the "MDS_L-layer".

The design principles of the data randomization part are rather clear. The principles behind the constructions in the key-schedule are not as easy to understand. However, we found no flaws in the description nor any reason to believe that the key-schedule is weak in any way.

2 Differential and linear cryptanalysis

In the following we evaluate Hierocrypt-3 with respect to differential and linear cryptanalysis. A difference of two bit-strings of equal lengths is defined via the exclusive-or operation.

The S-box. The S-box in Hierocrypt-3 is constructed from a power polynomial over $GF(2^8)$ together with an affine mappings used to destroy the mathematical structure. The power polynomial is $f(x) = x^{247}$ in $GF(2^8)$. It is well-known that an S-box constructed from an inverse function in a Galois field has highest achievable nonlinearity plus that the differential properties of this function are the best one can hope for. $f(x)$ has an easy connection to the inverse function. Note that $x^{247} = x^{-8} = (x^{-1})^{2^3}$ in $GF(2^8)$. Since squaring over $GF(2^8)$ is a linear function, it follows that $f(x) = x^{247}$ has differential and linear properties similar to that of x^{-1}.

The number of active S-boxes in a characteristic for Hierocrypt-3 are very easy to calculate for two rounds. In two consecutive rounds there will be at least 5 active 32-bit S-boxes (XS). Inside each active XS-box there will be at

least 5 active 8-bit S-boxes (S). In total for two rounds of Hierocrypt-3 there will be at least 25 active 8-bit S-boxes.

First we examine the subcomponents of Hierocrypt-3. It follows from the use of parity-check matrices of MDS-codes, that for MDS_L the branch number is five, that is, if there is s active S-boxes in the four-byte inputs, $s = 1, \ldots, 4$, then there are at least $5 - s$ active S-boxes in the four-byte outputs. And for MDS_H the branch number is also five, that is, if there is s active 32-bit XS-boxes in the inputs, $s = 1, \ldots, 4$, then there are at least $5 - s$ active XS-boxes in the outputs.

The S-box has a maximum differential and linear probability of 2^{-6} [27], see Appendix for the definition of linear probability.

Thus, in a traditional differential attack or in a traditional linear attack, the probabilities of two-round characteristics can be bounded by $(2^{-6})^{25} = 2^{-150}$. This means that for 4 rounds of Hierocrypt-3 the chances that a differential or linear attack will be applicable are very small. First of all, it would require that the attacker count over unrealisticly many key bytes, secondly, the probability of the involved characteristics are too small to facilitate any realistic attack.

3 Truncated differentials and linear hulls

The above analysis considered characteristics, see Appendix. Even stronger tools are differentials and truncated differentials. In the following we consider truncated differentials, but often refer to them as simply, differentials. It is possible that there exist several characteristics which can be combined into a differential. There have been block cipher cases in the past, where the differentials have a much higher probability than for corresponding characteristics. Also, it could be that several differential can be combined into a truncated differential.

For Hierocrypt-3 it seems that the best strategy for truncated differentials is to consider the values of blocks of 8 or 32 bits. That is, one only distinguishes between cases where blocks of 8 or 32 bits have a zero difference or a nonzero difference. However, note that the probabilities of truncated differentials must, in general, be higher than 2^{-128} to make any sense in an attack. The reason for this is, that the differential only specifies the exact values in a subset of all 128 bits, whereas the remaining bits can take any values. As an example, consider a truncated differential for Hierocrypt-3 which specifies only 32 bits. Then there are 96 bits in the differential which are not predicted. Thus for a randomly chosen permutation there is such a differential with a probability of 2^{-32}. If the truncated differential specifies say 64 bits, the probability should be higher than 2^{-64} to be distinguished from that of a randomly chosen permutation.

Assume first that a pair of 128-bit texts is specified by the exclusive-or difference, (x_0, x_2, x_2, x_3), where each x_i represents a 32-bit quantity. We shall write

$$(x_0, x_1, x_2, x_3) \xrightarrow{G} (y_0, y_1, y_2, y_3)$$

if texts of differences (x_0, x_1, x_2, x_3) can result texts of differences (y_0, y_1, y_2, y_3) after one application of a function G. Then the following type of differential could be possible

$$\begin{array}{ccc} (a,b,c,d) & \stackrel{XS}{\rightarrow} & (e,f,g,h) \\ (e,f,g,h) & \stackrel{MDS_H}{\rightarrow} & (i,0,0,0) \\ (i,0,0,0) & \stackrel{XS}{\rightarrow} & (j,0,0,0) \\ (j,0,0,0) & \stackrel{MDS_H}{\rightarrow} & (k,l,m,n) \end{array}$$

where it holds that a, b, c, and d and e, f, g, and h have values such that some of the byte differences involved are zeros. And similarly for k, l, m, n. This differential could be iterated to any number of rounds. Let us first ignore the probability through the first XS-layer. The probability of the above combinations through the MDS_H-layer is not known to us. It seems hard to calculate and the design principles behind MDS_H are not clear on this point. However, we conjecture that this probability is at most 2^{-48}, since it seems that this would involve at least six bytes in the differences to cancel out in the function MDS_H, but the probability is perhaps even lower. If the above differential is iterated to four rounds plus an XS-round, the bound on the probabilities is 2^{-96}. Since this differential would only specify the exact value of only 96 bits, this gives no advantage for an attacker as compared to attacking a randomly chosen permutation. Therefore, it seems that Hierocrypt-3 with five or more rounds is not vulnerable to an attack based on truncated differentials.

The above estimates can be translated into the case of linear hulls in much the same manner. Thus, we conclude that Hierocrypt-3 with five or more rounds does not seem vulnerable to an attack based on linear hulls.

4 Integral cryptanalysis

The best known attack on Rijndael [10] is an attack which was originally first applied to the block cipher Square [9]. The attack is generalised under the name of "integral cryptanalysis" in [21]. Since Hierocrypt-3 is reminiscent of Rijndael and uses very similar components we shall examine Hierocrypt-3 with respect to this attack.

In [2] the designers examine Hierocrypt-3 with respect to the integral attack and conclude that it applies to fewer S-box layers than Rijndael. However, as we shall show next there is an integral attack which applies to at least as many S-box layers as in the known attack on Rijndael.

Consider a collection of 2^{32} plaintexts which differ in one of the four 32-bit words, that is, the texts are all different in the inputs to one XS S-box but identical in the remaining three words. Then since XS is a bijective mapping this is also the case after the first XS-layer.

By construction of the MDS_H transformation, it holds that after the first MDS_H-layer, the texts are all different in each of the four 32-bit words. In other

words, in each word the 2^{32} texts take a value once only. But then since XS is a bijective mapping, it holds that after the second XS-layer, in each word the 2^{32} texts take a value once only. After the second MDS_H-layer the exclusive-or of all 2^{32} texts is zero in each of the four words. This follows by construction of the MDS_H linear transformation.

In summary there is an integral consisting of 2^{32} chosen plaintexts, for which the exclusive-or sum after two rounds of Hierocrypt-3 is zero in each of four 32-bit words. This integral can be used to attack Hierocrypt-3 reduced to five S-box layers. Note that two rounds of Hierocrypt-3 consists of four S-box layers. In the following when talking about "layers" we shall mean S-box layers. This is also the convention used by the designers.

Consider five layers of Hierocrypt-3 and the above integral. It follows that one can find the keys of the fifth layer, byte by byte, by guessing the value of one key byte and decrypt all texts through one layer of S-boxes. The values of the keys for which this sum is zero are candidates for the secret key. However, it is not necessary for each key byte to compute these values for all 2^{32} texts. It sufficies to consider the ciphertexts whose values in the particular byte occurs an odd number of times. The reason for this is, that if a value in one byte of the ciphertexts occurs an even number of times, then for any guess key byte, the exclusive-or sum of the values obtained by computing back through the S-box will sum to zero. Thus for each byte position one needs to consider only on the average 128 texts. To filter out any wrong value of the key, the above attack must be executed twice. In total one key byte in the last round can be determined in time about 2^{16} S-box evaluations.

All in all, there is an attack on Hierocrypt-3 reduced to five layers, which finds the last round key using about 2^{33} chosen plaintexts and in time equivalent to the time of doing 2^{19} S-box evaluations, which in time is equivalent to about 2^{12} encryptions. This should be compared to the designers' best attack on the same variant, which requires 2^{32} chosen texts and operates in time 2^{168}.

The attack can be extended to Hierocrypt-3 reduced to six layers by guessing an additional 32 bits of key material. This is similar to the attacks on Rijndael and to the attacks described by the designers [2]. In a first variant of this attack on Square [8] and on Rijndael [10] the time complexity was estimated to the time of about 2^{72} encryptions. However, in [12] this was improved to just 2^{44} encryptions by introducing short-cut strategies in the key-search. This improvement applies as well to the attack presented here on Hierocrypt-3. In summary, there is an attack on Hierocrypt-3 reduced to six layers, which finds the last round key using about 2^{36} chosen plaintexts and in time equivalent to the time of doing 2^{50} S-box evaluations, which in time is equivalent to about 2^{44} encryptions. This should be compared to the designers' own estimates that the integral attack is not applicable to Hierocrypt-3 reduced to six or more layers.

The attack on 6 layers can be extended to 7 layers by guessing even more key material, in a manner quite similar to the attacks on Rijndael. Thus, for Hierocrypt-3 with seven rounds, a simple extension the 6-round attack would require at least 2^{172} operations, so such attacks are clearly extremely expensive.

We conclude this section by noting that the integral attack applies at least

as well to six layers of Hierocrypt-3 as to six rounds of Rijndael. Also, we feel that there are some possibilities in trying to extend this to seven layers other than the above mentioned one for Rijndael. We are convinced however that the above method of integrals will not apply for the minimum twelve layers of Hierocrypt-3.

5 Other cryptanalysis

In this section we consider other attacks. 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 Hierocrypt-3, 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 nonlinear components of a low algebraic degree. Hierocrypt-3 uses S-boxes of a high nonlinear order and together with the relatively complex linear transformations, the probability that a higher order differential attack could be applicable is very small.

The slide attacks, the non-surjective attacks and the "mod n" attacks do not seem applicable to Hierocrypt-3 .

The interpolation attacks apply to ciphers which use simple mathematical functions only. Hierocrypt-3 uses mathematical functions in the S-boxes, however the affine mappings in the S-boxes together with good linear transformations have a good effect in thwarting the interpolation attacks.

The key-schedule of Hierocrypt-3 uses components of the data randomization part plus several linear transformations which together seem to avoid any clear patterns in the deduced round keys. We found no reasons to believe that there exist related-key attacks nor any (particularly) weak keys.

A last small remark is, that in §3.2.5 of [1] an operation ρ_0 is defined by removing $P^{(32)}$ from the operation ρ. We think the designers mean to define the function σ_0 from the function σ.

6 Survey of previous results

The only previous results on Hierocrypt-3 that we are aware of are those of the designers themselves [2].

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 [29] after the encryption of $2^{m/2}$ blocks, equal ciphertext blocks can be expected and information is leaked about the plaintexts [7, 17, 26].

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

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

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 [19]:

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 [20] and later to Skipjack [4]. 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 [22, 19] 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 $f : GF(2^n) \rightarrow GF(2^n)$ 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 $f : GF(2^n) \rightarrow GF(2^n)$ 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 [32] 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 [23]. A preliminary version of the attack on FEAL was described in 1992 [25]. Linear cryptanalysis [23] 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

$$(P \cdot \alpha) \oplus (C \cdot \beta) = (K \cdot \gamma) \tag{2}$$

which holds with probability $p_L \neq \frac{1}{2}$ over all keys [23], such that $|p_L - \frac{1}{2}|$, called the bias, is maximal. In (2) $P, C, \alpha, \beta, \gamma$ 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 [23] 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 [28].

Finally, in [24] 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 [15] 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 [16], reducing an exhaustive key search by almost a factor of four. The concept "related-key attack" was introduced by Biham [3], 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 [18] and Kelsey, Schneier, and Wagner [14] 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 [14, 11].

A.10 Interpolation attack

In [13] 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

$$f(x) = \sum_{i=1}^{n} y_i \prod_{1 \le j \le n, j \ne i} \frac{x - x_j}{x_i - x_j}. \tag{3}$$

$f(x)$ 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 (3) is known as the *Lagrange interpolation formula* (see e.g.,[6, 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 [31] 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 [5] the "slide attacks" were introduced, based on earlier work in [3, 16]. 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 [9, 8]. The attack on Square slightly modified also applies to the block ciphers Crypton and Rijndael [10].

In [21] 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] Toshiba Corporation. Specification on a Block Cipher: Hierocrypt-3. September 15, 2000.

[2] Toshiba Corporation. Self Evaluation: Hierocrypt-3. September 15, 2000.

[3] 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.

[4] 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.

[5] 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.

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

[7] 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.

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

[9] 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.

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

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

[12] N. Ferguson, J. Kelsey, B. Schneier, M. Stay, D. Wagner, and D. Whiting. Improved cryptanalysis of Rijndael. In B. Schneier, editor, *Fast Software Encryption, Seventh International Workshop*. Springer Verlag, 2001. To appear.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

[19] 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.

[20] 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.

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

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] K. Nyberg. Differentially uniform mappings for cryptography. In T. Helleseth, editor, *Advances in Cryptology - EUROCRYPT'93, LNCS 765*, pages 55–64. Springer Verlag, 1993.

[28] 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.

[29] 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.

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

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

[32] 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.