Analysis of MARS

January 12, 2001

Executive Summary

This report presents the results of a limited evaluation of the block cipher MARS.

No important weaknesses or flaws were found on MARS. The round function of MARS looks simple but is relatively complex to analyse because of the different natures of the involved components.

MARS is an iterated cipher which runs in 32 rounds, but the rounds are not of the same type, which is somewhat unusual. The middle 16 rounds are the conjectured cryptographically strong part of MARS. The outer forward eight rounds and backward eight rounds are faster and not keyed, and were introduced to make attacks in the inner rounds more complicated. We believe that with respect to the state-of-the-art a cryptanalytic attack on 16 middle (core) rounds of MARS alone is likely to be of a very high complexity.

Finally we mention that this report is the result of a limited time of review. A concentrated long analysis might reveal properties of MARS, which we were not able to detect, however it is felt with respect to the currently known attacks on block ciphers, the security margin of MARS is sufficiently high for many years.

Contents

1 Structural features and characteristics

MARS is an iterated block cipher with 128-bit blocks and allows for three different key sizes to be compliant with the AES [38].

MARS is a 32-round Feistel-like network, but where the round function is not identical in each round. There are two types of rounds in MARS, the so-called "wrapper rounds" and the so-called "core-rounds". The wrapper rounds are unkeyed and designed to make a quick scrambling of the data, the core rounds are keyed and supposedly the cryptographically strong rounds. First a 128-bit key is added word-wise modulo 2^{32} to the 128-bit plaintext. This key is also called the "pre-whitening key". Then the text is split into 4 words, each of 32 bits. The words are then input to eight wrapper rounds, also called the "forward mixing phase". The resulting words are then input to eight core-rounds, also called the keyed forward transformation. The resulting words are then input to another eight core-rounds, also called the keyed backward transformation. Finally the words are input to eight wrapper rounds, called the "backward mixing phase" and a final subkey is subtracted modulo 2^{32} to each word to form the outputs. The final key is also called the "post-whitening key".

The idea behind the forward and backward phases is that this structure makes the encryption operation similar to the decryption operation.

2 Differential cryptanalysis

In this section we evaluate MARS with respect to differential cryptanalysis. First we consider the cases where a difference of two bit-strings of equal lengths is defined via the exclusive-or operation. We shall examine the different components of the MARS rounds with respect to differential cryptanalysis.

The notation used is

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

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, where each x_i and y_i are 32-bit values.

Consider first the forward mixing phase, which consists of eight wrapper rounds. In each round there are four table lookups, using two S-boxes each taking an eight-bit input and returning a 32-bit output. It is not clear where and how to split the four wrapper rounds into four single wrapper rounds, but let us make the convention that a single wrapper round starts with an application of the S-box $S0$, and such that each wrapper round consists of two applications of the S-box $S0$ and two applications of the S-box $S1$. Consider two inputs each of 128 bits, such that these differ only in the most signficant byte of the most significant 32-bit word. Let i_{fmr} denote i forward mixing rounds. Then it holds that

$$(x, 0, 0, 0) \xrightarrow{4_{fmr}} (y, w, 0, z)$$

where x and z are zero in the least signficant 24 bits, y is zero in the most significant 24 bits, and w is a value not predicted. In particular, if x and z are

both set only in the most significant bit, and y is set only in the most significant bit of the least significant byte. The difference w in the second words is one of 256 possible nonzero 32-bit values. In plain words, this differential says that for two 128-bit texts different in only the most significant bit will after the first four wrapper rounds be equal in one of the four words, different only in one bit in each of two words, and the difference in the last words will be one of 256 predetermined differences. We shall continue this analysis with this specific value of x. Let us consider five wrapping rounds. It holds that

$$(x,0,0,0) \overset{4_{fmr}}{\rightarrow} (y,w,0,z) \overset{1_{fmr}}{\rightarrow} (s,t,0,u),$$

where t is closely related to w and s closely related to z. The value of t is closely related to w since the only thing that has happened to the texts of difference w is that some constant words have been added modulo 2^{32} to each texts in the difference. Also, s is closely related to z, since the four words before the very last operation in the fifth round differ in only the most significant bit of the least signficant bit, and in the last operation have been added modulo 2^{32} texts of (exor-)difference s. After the sixth wrapper round all four words have been affected by a differential:

$$(x,0,0,0) \overset{5_{fmr}}{\rightarrow} (s,t,0,u) \overset{1_{fmr}}{\rightarrow} (s',t',v,u),$$

but where in a similar manner as above, s' is related to s and t' is related to t. After seven wrapper rounds, the difference in the second words is the same as after six rounds, but all other words have been affected. These (strong) relations are not present after an additional round, in total eight wrapper rounds.

Now, let i_{bmr} denote i backward mixing rounds. Let x be a 32-bit difference set only in the most significant bit. Then the following holds

$$(x,0,0,0) \overset{4_{bmr}}{\rightarrow} (y,w,0,0),$$

where y is a difference with a 1-bit set only in the most significant bit of the least signficant byte. The difference w in the second words is one of 256 possible nonzero 32-bit values.

For five backward mixing rounds the following holds

$$(x,0,0,0) \overset{4_{bmr}}{\rightarrow} (y,w,0,0) \overset{1_{bmr}}{\rightarrow} (y',w,0,0),$$

where y' and y are related. Note that the Hamming weight of the difference y is one. The texts of difference y are added modulo 2^{32} a constant value plus exclusive-ored with a constant value. It holds [27] that with probability $1/2$ $y' = y$, with probability $1/4$ y' has a Hamming weight of two etc.

For six backward mixing rounds the following holds

$$(x,0,0,0) \overset{5_{bmr}}{\rightarrow} (y',w,0,0) \overset{1_{bmr}}{\rightarrow} (y',w',0,0),$$

where w' and w are related.

For seven and eight backward mixing rounds none of the four values can be predicted with certainty.

This analysis of the forward and backward mixing rounds shows that although none of the four words after eight forward or eight backward wrapper rounds can be predicted with certainty, the distribution of the differences in the four words is not uniform. Therefore 8 forward mixing rounds or 8 backward mixing rounds by themselves are not able to prevent high probability differentials. However in combination and assuming they were keyed (e.g., by adding key material to the input of each S-box) the sixteen rounds together would very likely well resist attacks based on differential cryptanalysis. This illustrates that the mixing rounds are important in helping to resist a differential attack on MARS.

Let us next consider the core rounds of MARS. The different components are fixed rotations, data-dependent rotations, the multiplication of an odd key modulo 2^{32}, the addition modulo 2^{32} and an S-box taking nine inputs bits and producing a 32-bit output. The core rounds work on 32-bit words. In one rounds, one of the 32-bits word is used as input to the round function E and three 32-bit words are output, which are added modulo 2^{32}, added modulo 2^{32}, and exored respectively to the three other input words of the round. After one round, the four 32-bit words are shifted one position, such that a different word is input to the function E in each of four (consecutive) rounds. E takes as input also two subkeys, where one of the keys, the one used in the multiplication operation, is always odd (when viewed as a 32-bit integer). In the first output word, the input is computed by first rotating the words by 13 positions to the left, the multiplying the odd subkey, then rotating the result by 10 positions to the left. The second output is computed by first adding the second subkey modulo 2^{32}, and then rotating the 32-bit result by an amount from an intermediate value from the computation of the first output. The third output is computed by first adding the second subkey modulo 2^{32}, then the least significant nine bits are input to the S-box, whose output is exclusive-ored by an intermediate value from the computation of the first output and then exclusive-ored by an intermediate value from the computation of the second output and finally the 32 bits are rotated by an amount from an intermediate value from the computation of the first output.

Let us first consider the components of the round function E. Let $\otimes_K$ denote a multiplication of a key modulo 2^{32} used in MARS. Then the following differentials hold with probability one

$$(a \mid \text{`0'}) \quad \stackrel{\otimes_K}{\rightarrow} \quad (A \mid \text{`0'}).$$

Here a and A denote some nonzero values of some t bits and "0" denotes a sequence of $32 - t$ bits. In other words, if a pair of texts are equal in the lower s bits $1 \leq s \leq 32$, then the texts after a multipication of a constant modulo 2^{32} have the same property. In particular, it holds with probability one, that two texts different in only the most significant bit have the same property after a

multiplication with an odd key.

The input to the S-box is 32 bits but only nine bits are used. Therefore there exist differentials of probability one through the application of an S-box, if the lower nine bits of the two 32-bit inputs are equal. When the lower nine inputs bits are different the differences in the outputs of the S-box are the exclusive-or of two (of totally 512) randomly chosen 32-bit quantities. Therefore, differentials which try to exploit the distribution of such differentials are expected not to be very effective in attacks on the full version of MARS.

A modular addition of a round key modulo 2^{32} has only a limited effect on differences of low Hamming weights. Let A and B be two 32-bit words which only differ in few bits. Then an integer addition of a (constant) key K does not necessarily lead to an increase of bit differences in the sums $A+K$ and $B+K$. This is illustrated in the following. Suppose the words A and B only differ in the most significant bit. Then it follows that $A+K$ and $B+K$ 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 $\frac{1}{2}$, $A+K$ and $B+K$ 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 K, the binary representation of the sum $Z = A + K$ may be obtained by the formulae

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

where σ_{j-1} denotes the carry bit and $\sigma_{-1} = 0$ (cf. [41]). Using these formulae one sees that $A+K$ and $B+K$ 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 $\frac{1}{4}$, $A+K$ and $B+K$ differ in exactly one bit and that with probability $\frac{3}{8}$, $A+K$ and $B+K$ differ in exactly two (not necessarily consecutive) bits. Thus with probability $\frac{5}{8}$ the words $A+K$ and $B+K$ 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 fixed keys can only moderately contribute to an avalanche effect of bit differences.

Next let us consider rotations. Using the exclusive-or operation to define a difference, it is clear that fixed rotations allow only for differentials of probability one. This does not mean that fixed rotations can be ignored, since the moving of certain bits to certain positions can have a dramatic effect on the differences in other components. Data-dependent rotations share some of these properties but only when the two texts in a differential are rotated by the same amount. Since the rotations are determined by the least significant five bits of a word, two words A and B different in one or several of these five bits, will have an difference after the rotations which depend on both A and B and not only on their difference (as is the case for fixed rotations).

Let us consider the function E. Consider the computation of the first output. From above it follows that texts which are equal in the lower bits but different in the higher bits can be well predicted after the multiplication operation. The

outputs of the multiplication operation are first rotated five positions to the left. This has the effect that the five most significant bits of the multiplication result are used as the amount by which the 32 bits in the computation of the second output are rotated. Thus, this rotation can be expected to depend on many on the 32 input bits to E. The intermediate first output is then rotated by another five positions to the left. Thus the amount used to rotate 32 bits in the computation of the third output are the sixth to tenth most significant bits of the output of the multiplication operation. Thus, these bits can also be expected to depend on many of the input bits to E. So, although differences in the higher bits only perform well through a multiplication operation, the E function is designed such that in these cases, this may have a dramatic effect on the differences in the outputs of the other two words. Consider the second output word. This is just the input to E added the second subkey, then rotated by some variable amount. Thus, a difference in two inputs to E of a low Hamming weight will result in the low Hamming weight difference in the second output words. However, note that in both the forward core rounds and in the backward core rounds, in the following round, this second output word does not become the input word to E. Before this word becomes the input to E, it has appeared in one other round and been added either a first or third output. Consider the third output word. If in a differential the lower nine bits are zeros, then the application of the S-box has no effect on the difference. The third output is exclusive-ored with 32 intermediate bits from the first output, then exclusive-ored with the final first output, and finally rotated by an amount determined from the final first output. Thus, if the differences in the inputs escape the S-box, then later there will be either different rotation amounts or the texts exclusive-ored from quantities from the computation of the first output word will be different and not easily predicted. Thus, it appears to be very difficult to predict the exact values in the differences of the outputs of E over several rounds with a high probability. This is supported by the report of the designers, which include crude estimates of the probabilities of differentials, plus by the results previously reported by other researchers. The best known results are those of [18, 19, 7], which all make us of a trivial 3-round differential, which is applied twice. The latter results will be reviewed later in this report.

We believe that a differential attack on the 16 core rounds by themselves is very unlikely to exist, and if it does it will have a very high complexity. Together with the relative good resistance of the 16 wrapper rounds in addition, it is safe to conclude that with respect to the state-of-the-art a differential attack on MARS is unlikely to exist.

3 Linear cryptanalysis

In this section we consider attacks based on linear cryptanalysis. In the following we examine the different components of the Feistel round function with respect to linear cryptanalysis.

Consider the 16 wrapper rounds and consider first the four output words

from four forward mixing rounds. It follows that each word is exclusive-ored or added modulo 2^{32} with the outputs of 4 S-box evaluations. In addition the third and fourth words are added intermediate values from the computations of the second and first words respectively. Consequently, a linear relation through eight mixing rounds needs to take many S-box linear relations into account. Since these are of low probabilities, such an approach is likely not to produce good results for neither eight forward mixing rounds nor for eight backward mixing rounds. Taking the designers own analysis for the sixteen core rounds into consideration which asserts that a linear attack is not possible, it is safe to conclude that for MARS as a whole a linear attack is very likely not to exist. What also speaks in favor of this claim is that linear cryptanalysis has proven most useful for ciphers which are limited to the exclusive-or operation and relatively small S-boxes.

4 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 MARS, 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. MARS uses S-boxes of a high nonlinear order in a relatively complex round function, and the probability that a higher order differential attack could be applicable is very small. Moreover, a dth order differential attack considers a collection of 2^d texts. The data-dependent rotations used in MARS should be equal for these 2^d texts in order that one should be able to predict the differential after the rotation. This makes the higher order differential attack very unlikely to succeed.

The slide attacks, the non-surjective attacks and the "mod n" attacks do not seem applicable, since the structure and components of MARS do not seem friendly to these attacks.

The integral attacks apply to MARS but only for a few of the core rounds, at most six rounds, we claim.

The interpolation attacks apply to ciphers which use simple mathematical functions only. The S-boxes used in MARS are generated in a pseudo-random fashion from the hash function SHA-1 [35]. This together with multiplications mod 2^{32}, the mixed use of exclusive-ors and modular additions and the different types of rounds make the interpolation attacks very unlikely to be applicable.

The key-schedule of MARS does not seem to allow for related-key attacks. The schedule is rather complex and involves S-box evaluations, fixed constants, rotations and both exclusive-ors and modular additions. Therefore, it seems unlikely that any easily identified weak keys exist.

5 Survey of previous results

In [42] it was shown that in the first proposed key-schedule of MARS it was possible to find so-called equivalent keys. That is, pairs of keys which produce the same set of round keys. This was possible also because that key-schedule allowed for keys up to 1248 bits. In [4] a new key-schedule was proposed, which is the current key-schedule of MARS. The before-mentioned equivalent keys do not exist for MARS with the new key-schedule.

In [9] it was shown that the MARS S-boxes do not satisfy exactly the criteria claimed by the designers. Also, in [29] it was shown that there exist linear relations in the 9 to 32 bit S-box of MARS higher than conjectured by the designers. However, none of these findings have been utilised in any improvement of cryptanalysis on MARS. Also, K. Aoki made an exhaustive search for all linear relations through this S-box and found many more examples than those of [29].

In [40] it is claimed that the linear analysis done by the designers of MARS was too optimistic. It was claimed that the bound on the best biases in a linear approximation on the core rounds was "only" 2^{-49} where the designers' bound was 2^{-69}. First of all, these numbers are only bounds, and do not represent the biases of linear approximations actually determined. It is therefore likely that any linear approximation will have a lower bias. Secondly, since a linear attack needs approximately b^{-2} texts to succeed where b is the bias, both numbers are low enough to conclude that MARS is not vulnerable to a linear attack taken into the account that the 16 mixing rounds also help prevent the success of a linear attack.

In [19] the authors consider a number of attacks on reduced-round variants of MARS. First the authors consider a MARS variant consisting of 8 forward mixing rounds, then 5 core rounds, and then finally 8 backward mixing rounds. First of all, attacks on this variant do not give a good picture of the strength of MARS itself. The attacker is guessing the values of the pre- and postwhitening keys, and therefore he can unwrap the wrapping rounds. These attacks require the time equivalent to more than 2^{230} encryptions, and at least 2^{197} bytes of memory. Also, the authors present an attack on a more realisticly downscaled version of MARS. The variant consists of 3 forward mixing rounds, then 3 forward core rounds, three backward core rounds, and finally 3 backward mixing rounds. The attack requires 2^{69} chosen plaintexts, 2^{73} bytes of memory, and the time equivalent to 2^{194} encryptions or more.

In [18] the authors state a so-called boomerang-amplifier differential attack, which is claimed to break MARS reduced to 11 core rounds. The attack requires 2^{65} chosen plaintexts, 2^{70} bytes of memory, and the time equivalent to 2^{225} encryptions or more. This differential attack is based on a 3-round differential of probability one, which is due to the fact that one word of a plaintext is not modified before the fourth round. The differential is used twice in a boomerang-fashion together with some specific, intrinsic properties of the E function.

In [7] the authors present an impossible differential on eight core rounds of MARS. This differential is based on the 3-round differential of probability

one which was discussed in the previous paragraph, together with some other intrinsic properties of the E function. It does not seem likely that this attack can be extended to all rounds of MARS not even to only the core rounds.

These attacks on MARS are completely unpractical and have no effect on the security on the unmodified MARS. One might even claim that the attacks show the strength of MARS.

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 [37] after the encryption of $2^{m/2}$ blocks, equal ciphertext blocks can be expected and information is leaked about the plaintexts [11, 23, 34].

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 [25]:

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 [26] and later to Skipjack [6]. 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 [30, 25] 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 [43] 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 [31]. A preliminary version of the attack on FEAL was described in 1992 [33]. Linear cryptanalysis [31] 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 [31], 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 [31] 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 [36].

Finally, in [32] 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 [21] 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 [22], reducing an exhaustive key search by almost a factor of four. The concept "related-key attack" was introduced by Biham [5], 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 [24] and Kelsey, Schneier, and Wagner [20] 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 [20, 15].

A.10 Interpolation attack

In [17] 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.,[10, 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 [39] 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 [5, 22]. 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 [13, 12]. The attack on Square slightly modified also applies to the block ciphers Crypton and Rijndael [14].

In [28] 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] IBM Corporation. MARS - a candidate cipher for AES.

[2] IBM Corporation. Self Evaluation Report.

[3] The IBM MARS team. Comments on MARS's linear analysis. May 12, 2000.

[4] The IBM MARS team. Suggested "tweaks" for the MARS cipher. Available at `www.nist.gov/aes`.

[5] 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.

[6] 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.

[7] E. Biham, V. Furham. Impossible Differentials on 8-Round MARS' Core. In proceedings of the Third AES Candidate Conference. Available at `www.nist.gov/aes`.

[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] L. Burnett, G. Carter, E. Dawson, W. Millan Efficient Methods for Generating MARS-like S-boxes. In proceedings of FSE'2000 to be published by Springer Verlag.

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

[11] 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.

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

[13] 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.

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

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

[16] T. Iwata, K. Kurosawa On the Pseudorandomness of AES Finalists – RC6, Serpent, MARS and Twofish. In proceedings of FSE'2000 to be published by Springer Verlag.

[17] 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.

[18] J. Kelsey, T. Kohno, B. Schneier. Amplified Boomerang Attacks Against Reduced-Round MARS and Serpent. In proceedings of FSE'2000 to be published by Springer Verlag.

[19] J. Kelsey, B. Schneier. MARS Attacks! Preliminary Cryptanalysis of Reduced-Round MARS Variants. In proceedings of the Third AES Candidate Conference. Available at `www.nist.gov/aes`.

[20] 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.

[21] 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.

[22] 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.

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

[24] 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.

[25] 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.

[26] 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.

[27] L.R. Knudsen and W. Meier. Differential cryptanalysis of RC5. *European Transactions on Telecommunications*, 8(5):445–454, September/October 1997.

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

[29] L. Knudsen, H. Raddum. Linear approximations to the MARS S-box. Available at `www.nist.gov/aes`.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] NIST. Secure hash standard. FIPS 180-1, US Department of Commerce, Washington D.C., April 1995.

[36] 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.

[37] 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.

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

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

[40] M. Robshaw, Y.L. Yin. Potential flaws in the conjectured resistance of MARS to linear cryptanalysis. Presented at the rump session of the Third AES Candidate Conference.

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

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

[43] 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.