Contemporary Block Ciphers

Lars R. Knudsen

University of Bergen, Department of Informatics, Hi-techcenter
N-5020 Bergen, Norway

Abstract. This paper considers modern secret-key block ciphers. The theory behind the design and analysis of modern block ciphers is explained, and the most important known attacks are outlined. Finally the Advanced Encryption Standard is discussed.

1 Block Ciphers - Introduction

In the last few thousands of years encryption algorithms, also called ciphers, have been developed and used [18,28]. Many of the old ciphers are much too weak to be used in applications today because of the tremendous progress in computer technology. There are essentially two types of cryptosystems, one-key and two-key ciphers. In one-key ciphers the encryption of a plaintext and the decryption of the corresponding ciphertext is performed using the same key. Until 1976 when Diffie and Hellman introduced *public-key* or two-key cryptography [20] all ciphers were one-key systems, today called conventional or classical cryptosystems. Conventional cryptosystems are widely used throughout the world today, and new systems are published frequently. There are two kinds of one-key ciphers, stream ciphers and block ciphers. In stream ciphers, typically a long sequence of bits is generated from a short string of key bits, and is then added bitwise modulo 2 to the plaintext to produce the ciphertext. In block ciphers the plaintext is divided into blocks of a fixed length, which are then encrypted into blocks of ciphertexts using the same key. The interested reader will find a comprehensive treatment of early cryptology in [28].

A block cipher is called an *iterated cipher* if the ciphertext is computed by iteratively applying a round function several times to the plaintext. In each round a round key is combined with the text input. In other words, let G be a function taking two arguments, such that, it is invertible when the first argument is fixed. Then define

$$C_i = G(K_i, C_{i-1}),$$

where C_0 is the plaintext, K_i is the ith round key, and C_r is the ciphertext. A special kind of iterated ciphers are the *Feistel* ciphers. A Feistel cipher with block size $2n$ and r rounds is defined as follows. Let C_0^L and C_0^R be the left and right halves of the plaintext, respectively, each of n bits. The round function G operates as follows

$$\begin{aligned} C_i^L &= C_{i-1}^R \\ C_i^R &= F(K_i, C_{i-1}^R) + C_{i-1}^L, \end{aligned}$$

I. Damgård (Ed.): Lectures on Data Security, LNCS 1561, pp. 105–126, 1999.

and the ciphertext is the concatenation of C_r^R and C_r^L. Note that F can be any function taking as arguments an n-bit text and a round key K_i and producing n bits. '+' is a commutative group operation on the set of n bit blocks. For the remainder of this paper we will assume that '+' is the exclusive-or operation ($\oplus$).

The Data Encryption Standard (DES) [55] is by far the most widely used iterated block cipher today. Around the world, governments, banks, and standards organisations have made the DES the basis of secure and authentic communication [65]. The DES is a Feistel cipher. However, the key size and the block size of the DES have become too small. Therefore the National Institute of Standards and Technology (NIST) in the U.S.A. has initiated the process of developing and to standardise a new encryption algorithm, the *Advanced Encryption Standard (AES)* [57], as a replacement for DES. This work is ongoing as this paper is written.

The remainder of this paper is organised as follows. § 2 lists and discusses the modes of operation for block ciphers used for encryption. § 3 discusses the theoretical and practical security of block ciphers. The most important methods of cryptanalysing block ciphers are given in § 4. § 5 discusses design principles of block ciphers and §6 reviews how to strengthen the DES. In §7 the Advanced Encryption Standard is discussed and some conjectures are made, and § 8 contains concluding remarks.

2 Modes of Operations

The most obvious and widespread use of a block cipher is for encryption. In 1980 a list of four modes of operation for the DES was published [56]. These four modes can be used with any block cipher and seem to cover most applications of block ciphers used for encryption [18]. In the following let $E_K(\cdot)$ be the permutation induced by using the block cipher E of block length n with the key K and let $P_1, P_2, ..., P_i, ...$ be the blocks of plaintexts to be encrypted. The Electronic Code Book (ECB) is the native mode, where one block at a time is encrypted independently of the encryptions of other blocks, $C_i = E_K(P_i)$, $P_i = E_K(C_i)$. In the Cipher Block Chaining (CBC) mode the encryption of a block depends on the encryptions of previous blocks. $C_i = E_K(P_i \oplus C_{i-1})$, $P_i = D_K(C_i) \oplus C_{i-1}$, where C_0 is a chosen initial value. The Cipher Feedback (CFB) mode is a stream cipher mode, where one m-bit character at a time is encrypted.

$$C_i = P_i \oplus \mathrm{MSB}_m(E_K(X_i))$$
$$X_{i+1} = \mathrm{LSB}_{n-m}(X_i) \parallel C_i$$

where X_1 is a chosen initial value, $\parallel$ denotes concatenation of blocks, MSB_s and LSB_s denote the s most and least significant bits respectively or equivalently the leftmost and rightmost bits respectively. Decryption is similar to encryption. Here m can be any number between 1 and the block length of the cipher. If the plaintext consists of characters, $m = 7$ or $m = 8$ is usually the well-chosen

parameter. The Output Feedback (OFB) mode is a second stream mode, where the stream bits are not dependent on the previous plaintexts, that is, only the stream bits are fed back, not the ciphertext as in CFB mode.

$$C_i = P_i \oplus \mathrm{MSB}_m(E_K(X_i))$$
$$X_{i+1} = \mathrm{LSB}_{n-m}(X_i) \parallel \mathrm{MSB}_m(E_K(X_i))$$

where X_1 is a chosen initial value. Decryption is equal to encryption. Both the CFB and OFB modes have two parameters, the size of the plaintext block and the size of the feedback value. In the above definition we have chosen them equal and will do so also in the following.

The ECB is the native mode, well-suited for encryption of keys of fixed length. It is not suited for the encryption of larger plaintexts, since equal blocks are encrypted into equal blocks. To avoid this, the CBC mode is recommended. Not only does a current ciphertext block depend on the current plaintext but also on all previous ciphertext blocks. In some applications there is a need for encryptions of characters, instead of whole blocks, e.g., the 8 bytes for the CBC mode of DES. For that purpose the CFB and OFB modes are suitable. It is often recommended to use the OFB mode only with full feedback, i.e., with $m = n$ (64 for the DES). It comes from the fact, that for $m < n$ the feedback function is not one-to-one, and therefore has a relatively short cycle [18] of length about $2^{n/2}$.

An important issue in the applications of the four modes is how an error in the transmission of ciphertexts is propagated. In the ECB mode an error in a ciphertext block affects only one plaintext block. A lost ciphertext block results in a lost plaintext block. An error in a ciphertext block in the CBC mode affects two plaintexts blocks. As an example, assume that ciphertext C_3 has an error and that all other ciphertext blocks are error-free, then $P_4 = D_K(C_4) \oplus C_3$ inherits the error from C_3 and $P_3 = E_K(C_3) \oplus C_2$ will be completely garbled. Here we assume that even a small change in the input to the block cipher will produce a randomly looking output. All other plaintexts will be decrypted correctly. A lost ciphertext block results in a lost plaintext block and an error in one addition plaintext block after which the mode synchronises itself. In the CFB mode an error in a ciphertext block C_i will be inherited by the corresponding plaintext block P_i, and moreover since X_{i+1} contains the garbled C_i the subsequent plaintexts blocks will be garbled until the X value is free of C_i, i.e., when C_i has been shifted out. In other words in CFB mode with m-bit ciphertexts, at most $n/m + 1$ plaintext blocks will be garbled. The case of lost ciphertext blocks is similar to that of the CBC mode. In the OFB mode, since the feedback is independent of the plaintexts and ciphertexts, a transmission error in a ciphertext block garbles only the corresponding plaintext block and is not propagated to other plaintext blocks. On the other hand, a lost ciphertext block will result in an infinite error propagation.

3 Security of Secret-Key Block Ciphers

When discussing the security of cryptographic systems one needs to define a model of the reality. We will use the model of Shannon [64]. The sender and the receiver share a common key K, which has been transmitted over a secure channel. The sender encrypts a plaintext P using the secret key K, sends C over an insecure channel to the receiver, who restores C into P using K. The attacker has access to the insecure channel and can intercept the ciphertexts (cryptograms) sent from the sender to the receiver. In this section we assume that the legitimate sender and receiver use a secret-key cipher $E_K(\cdot)$ of block size n (bits), where the key K is of size k. To avoid an attacker to speculate in how the legitimate parties have constructed their common key, the following assumption is made.

Assumption 1. *All keys are equally likely and a key K is always chosen uniformly random.*

Also we will assume that all details about the cryptographic algorithm used by the sender and receiver are known to the attacker, except for the secret key. This assumption is known as Kerckhoffs's Assumption [28].

Assumption 2. *The enemy cryptanalyst knows all details of the enciphering process and deciphering process except for the value of the secret key.*

For a fixed key, a block cipher is a permutation. There are totally 2^{n2^n} possible n-bit permutations. Thus, it would require $k = n2^n$ bits to specify all of them. With a block size of 64 bits or more this is a huge number. In a practical block cipher, the key size is much smaller, typically $k = 128$ or $k = 256$. A block cipher (system) with a k-bit key and blocks of n bits can be seen as an algorithm of how to select and specify 2^k of all 2^{n2^n} n-bit permutations.

3.1 Classification of Attacks

The possible attacks an attacker can do are classified as follows.

- *Ciphertext-only attack.* The attacker has obtained a set of intercepted ciphertexts.
- *Known plaintext attack.* The attacker obtains $P_1, P_2, ..., P_s$ a set of s plaintexts and the corresponding ciphertexts $C_1, C_2, ..., C_s$.
- *Chosen plaintext attack.* The attacker chooses *a priori* a set of s plaintexts $P_1, P_2, ..., P_s$ and obtains in some way the corresponding ciphertexts $C_1, C_2, ..., C_s$.
- *Adaptively chosen plaintext attack.* The attacker chooses a set of plaintexts $P_1, P_2, ..., P_s$ interactively as he obtains the corresponding ciphertexts $C_1, C_2, ..., C_s$. That is, the attacker chooses P_1, obtains C_1, **then** chooses P_2 etc.

– *Chosen ciphertext attacks.* For symmetric ciphers these are similar to those of chosen plaintext attack and adaptively chosen plaintext attack, where the roles of plain- and ciphertexts are interchanged.

Also, one can consider any combination of the above attacks. The chosen text attacks are obviously the most powerful attacks. In many applications they are however also unrealistic attacks. If the plaintext space contains redundancy, it will be hard for an attacker to 'trick' a legitimate sender into encrypting non-meaningful plaintexts and similarly hard to get ciphertexts decrypted, which do not yield meaningful plaintexts. But if a system is secure against an adaptively chosen plaintext/ciphertext attack then it is also secure against all other attacks. An ideal situation for a designer would be to prove that her system is secure against an adaptively chosen text attack, although an attacker may never be able to mount more than a ciphertext only attack.

3.2 Theoretical Secrecy

In his milestone paper from 1949 [64] Shannon defines perfect secrecy for secret-key systems and shows that they exist. Shannon's theory is described in many text books and here only a few of his results are stated. A secret-key cipher is *perfect* if for all P and all C it holds that $\Pr(P) = \Pr(P|C)$ [64]. In other words, a ciphertext C gives no information about the plaintext. This definition leads to the following result.

Corollary 1. *A perfect cipher is unconditionally secure against a ciphertext-only attack.*

As noted by Shannon the Vernam cipher, also called the *one-time pad*, is a perfect secret-key cipher. In the one-time pad the plaintext characters are exclusive-ored with independent key characters to produce the ciphertexts. However, the practical applications of perfect secret-key ciphers are limited, since it requires as many digits of secret key as there are digits to be enciphered [45]. Clearly, the above definition of a perfect cipher makes no sense when considering known or chosen plaintext attacks. A less stringent form of theoretical secrecy is possible, in terms of the *unicity distance*. It is the smallest integer s such that essentially only one value of the secret key K could have encrypted some plaintexts to the ciphertexts $C_1, ..., C_s$. The unicity distance depends on both the key size and on the redundancy in the plaintext space. Redundancy is an effect of the fact that certain plaintext characters appear more frequently than others. However, the unicity distance gives no indication of the computational difficulty in breaking a cipher, it is merely a *lower* bound on the amount of ciphertext blocks needed in a ciphertext-only attack. The concept of unicity distance can be adapted also to the known or chosen plaintext scenario. In these cases the redundancy of the plaintexts from the attacker's point of view is zero. Let k and n be the number of bits in the secret key respectively in the plaintexts and ciphertexts. If we assume that the keys are always chosen uniformly at random the unicity distance in a known or chosen plaintext attack is $\lceil k/n \rceil$.

3.3 Practical Secrecy

In the recent years cryptanalysis has been focused on finding the key K of a secret-key cipher. However, there are other serious attacks, which do not find the secret key. In the sequel Assumption 1 is used.

- *Total break.* An attacker finds the secret key K.
- *Global deduction.* An attacker finds an algorithm A, functionally equivalent to $E_K(\cdot)$ (or $D_K(\cdot)$) without knowing the key K.
- *Instance (local) deduction.* An attacker finds the plaintext (ciphertext) of an intercepted ciphertext (plaintext), which he did not obtain from the legitimate sender.
- *Information deduction.* An attacker gains some (Shannon) information about the secret key, the plaintexts or the ciphertexts, which he did not get directly from the sender and which he did not have before the attack.
- *Distinguishing algorithm.* An attacker is able to tell whether the attacked cipher is a randomly chosen permutation or one of the 2^k permutations specified by the secret key.

Clearly, this classification is hierarchical, that is, if a total break is possible, then a global deduction is possible and so on.

A global deduction is possible when a block cipher contains a "block structure". If certain subsets of the ciphertext are independent of certain subsets of the plaintext, then no matter how long the key is, the block cipher is vulnerable to a global deduction in a known plaintext attack. Also, in iterated block ciphers the round keys are sometimes generated in a one-way fashion [62,63,15,16]. So in attacks, which find the round keys, it may be impossible for the attacker to derive the actual value of the secret key, but at the same time the round keys enable the attacker to encrypt and decrypt. An instance deduction can be as dangerous as a total break, if the number of likely plaintexts is small. Consider the situation where the block cipher is used for encrypting a key in a key-exchange protocol. Here only one plaintext is encrypted and a total break is equal to an instance deduction. If the plaintext space is highly redundant an information deduction can be a serious problem. In general, the legitimate parties are often interested in that no information at all about the plaintexts and keys are obtained by any enemies. A distinguishing algorithm is the least serious attack. Let A be an attack (a distinguisher), which has access to a black box which is able to compute $E_K(\cdot)$ for K the secret key. When asked for the ciphertexts of plaintexts $P_1, \ldots, P_i$ the black box flips a coin whether to return $E_K(P_1), \ldots, E_K(P_i)$ or $\pi(P_1), \ldots, \pi(P_i)$ for a randomly chosen permutation π. The attack A has to decide whether the encryptions came from $E_K(\cdot)$ or π. The advantage of the attack is $\text{abs}(\Pr(A : \text{"it is } E_K(\cdot)\text{"} | E_K(\cdot) \text{ was used}) - \Pr(A : \text{"it is } E_K(\cdot)\text{"} | \pi \text{ was used}))$, that is, a number between 0 and 1. The higher the number the better the attacker's strategy.

In the following some trivial attacks applicable to all block ciphers are discussed. All block ciphers are totally breakable in a ciphertext-only attack, simply by trying all keys one by one and checking whether the computed plaintext is

meaningful, using only about N_{ud} ciphertext blocks, where N_{ud} is the unicity distance. This attack requires the computation of about 2^k encryptions. Also, there is the table look-up attack, where the attacker encrypts in a pre-computation phase a fixed plaintext P under all possible keys and sorts and stores all the ciphertexts. Thereafter the cipher is total breakable in a chosen plaintext attack requiring one chosen plaintext. There might be some keys encrypting P into the same ciphertext. Repeating the attack a few times with $P' \neq P$ will give a unique key. All block ciphers are globally/instance deducible under a known/chosen plaintext attack. Simply get and store all possible plaintext/ciphertext pairs. The running time of a deduction is the time of one table look-up.

The following result shows that a non-trivial information gain can be obtained when about the square root of all ciphertexts are available.

Theorem 1 ([34]). *Every n-bit block cipher used in the ECB, CBC or CFB mode is information deducible in a ciphertext-only attack with complexity about $2^{n/2}$.*

Note that the result of Theorem 1 is independent of the key size. This attack on CBC mode was named the *matching ciphertext attack* in [12]. Thus, it is recommended that a single key is used to encrypt at most $2^{n/2}$ ciphertext blocks.

Hellman [24] has presented a time-memory trade-off attack on any block cipher, which finds the secret key after $2^{2k/3}$ encryptions using $2^{2k/3}$ words of memory. The $2^{2k/3}$ words of memory are computed in a pre-processing phase, which takes the time of 2^k encryptions.

To estimate the complexity of a cryptanalytic attack one must consider at least the time it takes, the amount of data that is needed and the storage requirements. For an n-bit block cipher the following complexities should be considered. *Data complexity:* The amount of data needed as input to an attack. Units are measured in blocks of length n. Denote this complexity C_d. *Processing complexity:* The time needed to perform an attack. Time units are measured as the number of encryptions an attacker has to do himself. Denote this complexity C_p. *Storage complexity:* The words of memory needed to do the attack. Units are measured in blocks of length n. Denote this complexity C_s. As a rule of thumb, the complexity of an attack is taken to be the maximum of the three complexities, that is, $C_a = max(C_d, C_p, C_s)$. In general, there are some deviations from this rule and furthermore the three complexities are relative to the attacker. As an example, we may say that the above attack by Hellman on the DES has complexity $2^{2\times 56/3} \simeq 2^{38}$. Although the time of the pre-computation phase is 2^{56}, it is done only once after which any DES-key can be derived with a complexity of 2^{38}. On the other hand, the storage requirements may be unrealistic for most attackers, e.g., the attack on the DES will require about 1000 Gigabytes of memory.

4 Cryptanalysis of Block Ciphers

The history of cryptanalysis is long and at least as fascinating as the history of cryptography. As a single example, in 1917 in an article in "Scientific American"

the Vigenère cipher was claimed to be "impossible of translation" [19]. Today, it is an exercise in cryptology classes to illustrate that this claim is not true.

4.1 Attacks on Iterated Ciphers

In the following, P denotes the plaintext and C denotes the ciphertext. In most modern attacks on iterated ciphers, the attacker repeats his attack for all possible values of (a subset of) the bits in the last-round key. The idea is, that when he guesses the correct values of the bits of the key, he can compute bits of the ciphertexts after the second-last round, that is before the last round, whereas when he guesses wrongly, these bits will correspond to ciphertext bits encrypted with a wrong key. If there is a probabilistic correlation between the bits of the plaintexts, P, and the bits of the ciphertexts before the last round, $\tilde{C}$, denoted $\mathrm{cor}(P, \tilde{C})$, an attacker might be able to distinguish the correct guesses of the key in the last round from wrong guesses. If this is the case, the attacker can peel off one round of the cipher and do a similar attack on a cipher one round shorter to find the second-last round key etc. In some attacks it is advantageous to consider the first-round key instead of the last-round key or both at the same time, depending on the structure of the cipher, the number of key bits involved in each round etc. In iterated ciphers the correlation is often found by first identifying a correlation between inputs and outputs of the individual rounds and them combining them to a correlation over several rounds. The probability of this correlation can be calculated as the product of the probabilities of the individual round correlations, if they are independent. For most ciphers this independence is obtained by assuming that all round keys are independent. Although this is most often not the case, first of all, experiments have shown [6,34,49] that this leads to a good approximation to the real probability, secondly there seems to be no other way to compute the real probability. Denote by the *reduced cipher*, the cipher that one gets by removing the first and/or the final rounds of the original cipher. Let $\tilde{P}, \tilde{C}$ be the input bits and output bits respectively of the reduced cipher. Let $\tilde{K}$ be the key bits the attacker guesses in the attack (note that an attacker might not need to know all input and output bits of the reduced cipher). If the attacker guesses $\tilde{K}$ correctly, he can compute (bits of) $\tilde{P}, \tilde{C}$ from P, C. Let P', C' be the results the attacker obtains when he guesses $\tilde{K}$ wrongly. The probability of success of an iterated attack depends first of all on whether $\mathrm{cor}(\tilde{P}, \tilde{C})$ is different from $\mathrm{cor}(P', C')$, at least for some wrong guesses of $\tilde{K}$. In most attacks on iterated ciphers, an attacker repeats the basic attack a number of times and counts the values of $\tilde{K}$ which led to the expected $\mathrm{cor}(\tilde{P}, \tilde{C})$. Although some attacks in the literature do not have exactly this form, they can be translated into this general form (at least for illustration). A similar approach was taken in [67]. The signal-to-noise ratio (see [6] for the differential attack) is the expected number of times the correct guess of the key is counted over the expected number of times a wrong guess of the key is counted. Earlier it was believed that a necessary condition for the success of an iterated attack is that the signal-to-noise ratio is greater than one [6]. However, it was later discovered [60,9] that an attack can work in two ways: when $S/N > 1$ one looks for the

most suggested value of the key, and when $S/N < 1$ one looks for the least suggested value. Attacks where $S/N < 1$ are in principle as good as attacks where $S/N > 1$ but do not seem easier to find in general. In the following a number of iterated attacks are described. Since all of them have the above form, it suffices to describe how to detect and obtain the correlation of bits of the inputs and outputs of the reduced cipher.

4.2 Differential Cryptanalysis

The most general method of analysing conventional cryptosystems today is *differential cryptanalysis*, published by Biham and Shamir in 1990. The method has proved to be relatively efficient and has been applied to a wide range of iterated ciphers see e.g., [6,32]. Furthermore, it was the first attack which could (theoretically) recover DES keys in time less than the expected cost of exhaustive search [6,7]. In the following a brief description of differential cryptanalysis with respect to a general n-bit iterated cipher, cf., (1) is given.

First, 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, so the difference is independent of the key. In a strong encryption algorithm there will be some components which are non-linear in the $\otimes$-operation. 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 (([6])). *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 $1 \leq i \leq s$.*

Define p_i as the probability that inputs of difference α_{i-1} lead to output of difference α_i, where the probability is taken over all choices of the round key and the inputs to the ith round. In [44] the notion of a Markov cipher was introduced. In a Markov cipher this probability is independent of the actual inputs of the round and is calculated over all possible choices of the round key. Also in [44] it was shown that in a Markov cipher if the round keys K_i are independent, the p_i's are also independent and

$$\Pr(\Delta C_s = \alpha_s \,|\, \Delta P_0 = \alpha_0) = \prod_{i=1}^{s} \Pr(\Delta C_i = \alpha_i \,|\, \Delta C_{i-1} = \alpha_{i-1}). \tag{2}$$

In some differential attacks using an $(r-1)$-round characteristic only the plaintext difference ΔP and the last ciphertext difference ΔC_{r-1} need to be fixed. That is, the intermediate differences $\Delta C_1, \Delta C_2, \ldots, \Delta C_{r-2}$ can have any value. Lai and Massey introduced the notion of *differentials* [44].

Definition 2. *An s-round* differential *is a pair of differences $\{\alpha_0, \alpha_s\}$, where $\Delta P = \alpha_0$, $\Delta C_s = \alpha_s$.*

The probability of an s-round differential $(\Delta P, \Delta C_s)$ is the conditional probability that given an input difference ΔP at the first round, the output difference at the sth round will be ΔC_s. More formally, the probability of an s-round differential is given as

$$\Pr(\Delta C_s = \beta_s \mid \Delta P = \beta_0) = \sum_{\beta_1} \cdots \sum_{\beta_{s-1}} \prod_{i=1}^{s} \Pr(\Delta C_i = \beta_i \mid \Delta C_{i-1} = \beta_{i-1}), \tag{3}$$

where $\Delta C_0 = \Delta P$. A differential will, in general, have a higher probability than a corresponding characteristic. Differentials were used in [54] to construct cipher secure against differential attacks. Also, for some ciphers there is a significant advantage in considering differentials instead of characteristics [40].

In a differential attack the attacker does not know the key. Therefore in finding a good differential, the attacker computes the probabilities of differentials assuming that all the round keys are uniformly random and independent. However, the pairs of encryptions an attacker gets are encrypted using the same key, where the round keys are fixed and (can be) dependent. In [42] this problem is dealt with as follows

Definition 3 ((Hypothesis of stochastic equivalence)). *For virtually all high probability $(r-1)$-round differentials (α, β)*

$$Pr_P(\Delta C_1 = \beta \mid \Delta P = \alpha,\ K = k) \approx Pr_{P,K}(\Delta C_1 = \beta \mid \Delta P = \alpha,)$$

holds for a substantial fraction of the key values k.

In the differential attack on IDEA in [9], it was exploited that the hypothesis of stochastic equivalence does not hold for IDEA reduced to 3.5 rounds. A differential attack was mounted for which the S/N-ratio is one when the differential is averaged over all keys. When the key is fixed the S/N-ratio is different from one and the secret key can be recovered with sufficiently many pairs of plaintexts and ciphertexts. In [38] a differential attack on DEAL is described using a differential of probability zero. Also, recently a differential attack with $S/N < 1$ on Skipjack was announced [5].

Experiments have shown that the number of chosen plaintexts needed by the differential attack in general is approximately c/p, where p is the probability of the differential being used and c a small constant.

Higher Order Differentials In [43] a definition of higher order derivatives of discrete functions was given. Later higher order differentials were used to cryptanalyse ciphers presumably secure against conventional differential attacks [37]. In [27] these attacks were extended and applied to the cipher of [54]. A dth order differential is a collection of 2^d (first-order) differentials. The main idea in the

higher order differential attack is the fact that a dth order differential of a function of nonlinear order d is a constant. Consequently, a $d+1$st order differential of the function is zero. Assume that (a subset of) the output bits of the reduced cipher are expressible as a low-degree polynomial $p(\tilde{x}) \in GF(2)[\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_i}]$, where $\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_i}$ is a subset of input bits to the reduced cipher. If this polynomial has degree not higher than d, then $\sum_{\tilde{x} \in \mathcal{L}_d} p(\tilde{x}) = c$, where $\mathcal{L}_d$ denotes a d-dimensional subspace of $GF(2)^n$ and c a constant. This method was applied to the cipher example given in [54]. This cipher is "provably secure" against a differential attack but can be broken in a higher order differential attack with relatively low complexity.

Truncated Differentials In some ciphers it is possible and advantageous to predict the values of only parts of the differences after each round of the cipher. Let $\{\alpha_0, \alpha_1, \ldots, \alpha_s\}$, be an s-round *characteristic*. Then $\{\alpha'_0, \alpha'_1, \ldots, \alpha'_s\}$ is called a truncated characteristic, if α'_i is a subsequence of α_i. Truncated characteristics were used to some extent in [6] but only in the outer rounds of a cipher. Note that a truncated characteristic is a collection of characteristics and therefore reminiscent of a differential. A truncated characteristic contains all characteristics $\{\alpha''_0, \alpha''_1, \ldots, \alpha''_s\}$ for which $\text{trunc}(\alpha''_i) = \alpha'_i$, where $\text{trunc}(x)$ is the truncated value of x, where the truncation is not further specified here. The notion of truncated characteristics extends in a natural way to truncated differentials introduced in [37].

The truncated differentials were used in [39] to attack SAFER K [46,47]. Also, in [9] truncated differential attacks were presented on IDEA [44] and latest on Skipjack [5].

4.3 Linear Cryptanalysis

Linear cryptanalysis was proposed by Matsui in 1993 [48]. A preliminary version of the attack on FEAL was described in 1992 [51]. Linear cryptanalysis is a known plaintext attack in which the attacker exploits linear approximations of some bits of the plaintext and ciphertext. In the attack on 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

$$(\tilde{P} \cdot \alpha) = (\tilde{C} \cdot \beta) \tag{4}$$

where α, β are n-bit strings and where '$\cdot$' denotes the dot product, which holds with probability $p_L \neq \frac{1}{2}$ over all keys, such that $|p_L - \frac{1}{2}|$, called the bias, is maximal. As in differential cryptanalysis one can define characteristics to be used in linear cryptanalysis.

The number of known plaintexts needed such that the relation (4) can be effectively detected is approximately $|p_L - 1/2|^{-2}$. The following result appears in [53].

Theorem 2. *If X and K are independent and K is uniformly distributed, then for all $a \in GF(2)^m$, $b \in GF(2)^n \in GF(2)^\ell$*

$$2^{-\ell} \sum_{k \in GF(2)^\ell} |\, P_X(X \cdot a + Y(X,k) \cdot b = 0) - 1/2 \,|^2 = \sum_{c \in GF(2)^\ell} |\, P_{X,K}(X \cdot a + Y(X,K) \cdot b + K \cdot c = 0) - 1/2 \,|^2$$

This theorem shows the similarity between the concept of differentials in differential cryptanalysis and in linear cryptanalysis. An expression of the form (4) is called a *linear hull*. Note that in [48] the linear approximations have the form $(\tilde{P} \cdot \alpha) = (\tilde{C} \cdot \beta) \oplus (K \cdot \gamma)$, where $(K \cdot \gamma)$ is an exclusive-or of round-key bits accumulated in the linear characteristic. The bias of the linear approximations is taken as the bias of the linear characteristic used. However, such an attack cannot be guaranteed to work in general. If there exist linear approximations such that $(\tilde{P} \cdot \alpha) = (\tilde{C} \cdot \beta) \oplus (K \cdot \gamma)$, and $(\tilde{P} \cdot \alpha) = (\tilde{C} \cdot \beta) \oplus (K \cdot \gamma')$ both with probability $p > 1/2$ but where $(K \cdot \gamma) \neq (K \cdot \gamma')$, then these two linear approximations may cancel the effect of each other. This was also noted in [3].

In Matsui's attack on the DES, experiments indicate that the bias of the linear hull is equal to the bias of a single characteristic [49]. It is further confirmed by computer experiments that the probability of (4) is close to 1/2 when the value of $\tilde{K}$ is wrong. It is estimated that the complexity of a linear attack on the DES with up to 16 rounds is about

$$N_P \simeq c \times |p_L - 1/2|^{-2}$$

where $c \leq 8$. The attack on the DES was implemented in 1994, required a total of 2^{43} known plaintexts [49] and is today the fastest, known key-recovery attack on the DES.

In [29] an improved linear attack using multiple linear approximations was given. In [41] a linear attack is shown using non-linear approximations in the outer rounds of an iterated cipher. For the DES none of these attacks have yet shown to offer an significant improvement compared to Matsui's linear attack. The attacks seem best suited for attacks on ciphers with large S-boxes.

4.4 Davies' Attack

In [17] a correlation attack on the DES was outlined. It exploits that the outputs from neighbouring S-boxes are not uniformly distributed. The correlation can be iterated to any number of rounds with a corresponding decrease in the probability. The attack was improved in [4] and finds the secret key of the DES using about 2^{50} known plaintexts, and is the third, known key-recovery attack which finds the secret key faster than by an exhaustive search.

4.5 Differential-Linear Attack

In [25] it was shown how to combine the techniques of differential and linear attacks. The attack is a chosen plaintext attack and considers pairs of plaintexts and ciphertexts, the bits of which are (partly) approximated by linear approximations. In particular, an attack on the DES reduced to 8 rounds was devised, which on input only 512 chosen plaintexts finds the secret key. It seems that the attack is not easily extended to more than 8 rounds of DES [25]. In [1] the differential-linear attack was applied to FEAL. The attack takes a long time, but only 12 chosen plaintexts are needed.

4.6 Other Variants

Several generalisations of the differential and linear attacks have been developed. In [67] a generalisation of both the differential and linear attacks, known as *statistical cryptanalysis* was introduced. It was demonstrated that a statistical attack on the DES included the linear attack by Matsui but without any significant improvement. The applications to other ciphers have not been demonstrated. In [22,23] two generalisations of the linear attack were given. However, none of them have yet proved to be much more efficient than the linear attack.

4.7 Interpolation Attack

In [27] the interpolation attack was introduced 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

$$f(x) = \sum_{i=1}^{n} y_i \prod_{1 \le j \le n, j \ne i} \frac{x - x_j}{x_i - x_j}. \tag{5}$$

$f(x)$ 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 (5) is known as the *Lagrange interpolation formula* (see e.g., [10, page 185]).

In the interpolation attack an attacker constructs polynomials using inputs and outputs of the reduced cipher. This is particularly easy if the components in the cipher can be easily expressed as mathematical functions. The idea in 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. In an extended version of the attack meet-in-middle techniques are used to further reduce the degrees of the used polynomials [27].

Recently, a probabilistic version of the interpolation attack was introduced [26].

4.8 Non-surjective Attack

In [61] the non-surjective attack on iterated ciphers was described. It is applicable to Feistel ciphers where the round function is not surjective. In a Feistel cipher the plaintexts and corresponding ciphertexts give the exclusive-or of all outputs of the round function. Thus, if the round function is not surjective this gives information about intermediate values in the encryptions, which can be used in an attack.

4.9 Key Schedule Attacks

In this section we consider the key schedules of block ciphers. We consider an n-bit block cipher, where $E_K(\cdot)$ denotes encryption with the key K and $D_K(\cdot)$ denotes decryption. A weak key K, is a key for which encryption equals decryption, that is, $E_K(X) = D_K(X)$ for all n-bit texts X. A pair of semi-weak keys K, K^*, are keys for which encryption with one keys equals decryption with the other key, that is, $E_K(X) = D_{K^*}(X)$ for all n-bit texts X or equivalently, $D_K(X) = E_{K^*}(X)$ for all n-bit texts X. It is well-known that there are at least four weak keys and six pairs of semi-weak keys for the DES. In [11] it was shown that there are exactly 2^{32} fixed points for the DES used with a weak key.

If there are only a small number of weak keys they pose no problem for applications of encryption if the used keys are chosen uniformly at random. However, when block ciphers are used in hash modes where e.g., the key input can be chosen by the attacker in attempts to find collisions, they play an important role as demonstrated in [14,59].

[13] lists a large class of 2^{51} keys for IDEA, which can be easily identified using only a few plaintexts and ciphertexts. Note that IDEA uses 128-bit keys. In [68] it was shown that for 1 in 2^{15} keys for Blowfish a differential attack is faster than an exhaustive key search. [40] lists a large class of differentially weak keys for RC5 [62], keys for which a specific differential attack has improved performance.

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.

The first kind of attacks was introduced in [33], the second kind of attacks in [2]. Also, there are related key attacks on SAFER K [36] and on several other block ciphers [30].

Note that for the attacks of 2b above one must omit Assumption 1. 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 quite realistic settings, in which an attacker may succeed to obtain such encryptions, as argued in [30]. Also, there exists quite efficient methods to preclude the related key attacks [30,16].

5 Design of Block Ciphers

In this section we discuss some of the problems involved in the design of a block cipher. Two generally accepted design principles for practical ciphers are the principles of confusion and diffusion that were suggested by Shannon. Massey[45] interprets Shannon's concepts of confusion and diffusion [64] as follows *Confusion:* "The ciphertext statistics should depend on the plaintext statistics in a manner too complicated to be exploited by the cryptanalyst". *Diffusion:* "Each digit of the plaintext and each digit of the secret key should influence many digits of the ciphertext". These two design principles are very general and informal. Shannon also discusses two other more specific design principles. The first is to make the security of the system reducible to some known difficult problem. This principle has been used widely in the design of public-key systems, but not in secret-key ciphers. Shannon's second principle is to make the system secure against all known attacks, which is still the best known design principle for secret-key ciphers today.

There have been many suggestions in the past of more specific design principles, e.g. completeness, strict avalanche criterion, see [52, page 277-278]. However a specific cryptographic design principle should not be overvalued. Design principles should be seen as "guidelines" in the construction of ciphers, evolved from years of experience, and as necessary, but *not* sufficient requirements. There are many examples of this in the history of cryptography. We already mentioned the example of [27], where a block cipher "provably secure" against differential and linear attacks was broken by some other means.

5.1 Block and Key Size

It is clear from the discussion in Section 3.3 that if either the block or key size is too small or both, a block cipher is vulnerable to a brute force attack. These attacks are independent of the internal structure and intrinsic properties of an algorithm. Most block ciphers in use today have a block size of 64 bits. For these ciphers the birthday attacks of Theorem 1 require storage/collection of 2^{32} ciphertext blocks for a success of about one half. It may seem unlikely that a single key is used to process that many ciphertexts, and the storage of 2^{32} ciphertext blocks of each 64 bits will require about 2^5 Gigabytes of memory. However with the rapid increase in computing power and available storage media it can expected that in a few years this attack is very realistic. This has be taken into consideration in the ongoing development of the Advanced Encryption Standard, cf. later.

The key size of the DES is only 56 bits, which is too short. In [69,70] a design of an exhaustive search machine was given, which at the cost of 1 million US$

finds the secret key of the DES in average time 0.5 hours. In [8] is was estimated that with respect to an exhaustive key search a key size of at least 90 bits will suffice for the next 20 years.

5.2 Resistance against Differential and Linear Attacks

We consider an r-round iterated block cipher with round function G. Denote by p_d the highest probability of a non-trivial one-round differential achievable by the cryptanalyst. Let p be the probability of a linear approximation. Then $|p - 1/2|$ is called the bias. Recall that the success of a linear attack is proportional to the reciprocal value of the square of the bias of the used linear approximation. It has been shown how to treat differential and linear cryptanalysis in a similar way [50] by defining $q = (2p - 1)^2$. Let q_ℓ denote the highest such quantity for a one-round linear approximation. It is possible to lower bound the probability of any differential and any hull in an r-round iterated cipher expressed in terms of p_d and q_ℓ.

Theorem 3 ([34]). *Consider an r-round iterated cipher, which has independent round keys. Any s-round differential, $s \geq 1$, has a probability of at most p_d. Any s-round linear hull, $s \geq 1$, has a reciprocal squared bias of at most q_ℓ.*

For Feistel ciphers, Theorem 3 is trivial, since $p_d = q_\ell = 1$ when the right halves of the inputs are fixed. These differentials and hulls are called trivial one-round differentials and hulls. It is possible to lower bound the probabilities of differentials and hulls in a Feistel cipher expressed in terms of the most likely non-trivial one-round differential with probability p_{max} and the best non-trivial one-round linear hull with reciprocal squared bias of q_{max}.

Theorem 4 ([54,50]). *Consider an r-round Feistel cipher with independent round keys. Any s-round differential, $s \geq 4$, has a probability of at most $2p_{max}^2$. Any s-round linear hull, $s \geq 4$, has a reciprocal squared bias of at most $2q_{max}^2$.*

It has been shown that the round function in a Feistel cipher can be chosen in such a way that p_{max} and q_{max} are small [54,34].

5.3 Resistance against other Attacks

As mentioned earlier one should be careful not to focus too much on the resistance against a limited set of attacks, when constructing new block ciphers. In some cases other attacks become possible.

Let E be a n-bit r-round iterated block cipher. Assume that the nonlinear order of the ciphertext bits after one round is d and d^s after s rounds with a high probability. Then higher order differential attacks will in general not be possible after r rounds, if $d^r \simeq n$. One should take into account that the attacker may be able to guess key bits in the outer rounds of the cipher thereby attacking a cipher with a fewer number of rounds. Thus, if the nonlinear order should reach the block size after, say, $r - 2$ rounds.

It is yet unknown how to obtain exact security against truncated differential attacks. However, a truncated differential is a collection of differentials. Therefore, if the probabilities of all differentials can be bounded sufficiently low, this attack will have only small probability of succeeding.

The differential-linear attack will only work if both good linear hulls and good differentials exist. Thus, the techniques of the previous section also apply in this case.

The interpolation attack works particularly well when the outputs of one round of a cipher can be described as a polynomial of the input bits with relatively few nonzero coefficients. Thus, if a cipher consists of elements which cannot be described as such, it seems that the attack will not be possible. The probabilistic version of the interpolation attack might improve on this, but this has not been reported and needs further study.

The key-schedule attacks can be precluded by using only so-called strong key-schedules [35], see also [30,16].

6 Enhancing the Strength of the DES

Already in 1977 the DES was criticised for its short key length and it was suggested to use the DES in a triple encryption mode [21]. In a triple encryption with three independent keys K_1, K_2, and K_3, the ciphertext corresponding to P is $C = E_{K_3}(E_{K_2}(E_{K_1}(P)))$. One variant of this idea is well-known as two-key triple encryption, proposed in [66], where the ciphertext corresponding to P is $E_{K_1}(D_{K_2}(E_{K_1}(P)))$. Compatibility with a single encryption can be obtained by setting $K_1 = K_2$. However, whereas triple encryption is provably as secure as single encryption, a similar result is not known for two-key triple encryption. A two-key triple encryption scheme with a proof of security appeared in [16]. Another method of increasing the key size is DES-X, developed by Rivest. In DES-X the ciphertext corresponding to P is $C = E_K(P \oplus K_1) \oplus K_2$, where K is a 56-bit key, and K_1 and K_2 are 64-bit keys. Alternatively, $K_1 = K_2$ may be used. It was shown [31] that for attacks not exploiting the internal structure the effective key size of DES-X is $118 - \log_2 m$ bits, where m is the maximum number of plaintext/ciphertext pairs the attacker can obtain.

Although all these schemes increase the key lengths of the DES, the block lengths of 64 bits of these proposals are the same as for DES, and the matching ciphertext attack is still a problem.

7 The Advanced Encryption Standard

A better solution than those of the previous section seems to be to construct a new block cipher with larger keys and larger blocks to replace the DES, a cipher which at the same time is immune to all kinds of attacks reported so far in the cryptographic literature. Such an initiative was announced in January 1997 by the U.S. National Institute of Standards and Technology (NIST), the same institute that standardized DES in the 70's. The first workshop was held

April 15, 1997. NIST's intention is to standardize a new encryption algorithm, the *Advanced Encryption Standard (AES)* [57], as a replacement for DES. NIST encouraged parties world-wide to submit proposals for the new standard. Submission deadline was June 15, 1998; 15 proposals from all over the world were submitted and all proposals are now publicly available [58]. The proposals are required to support at least a block size of 128 bits, and three key sizes of 128, 192, and 256 bits. NIST hopes that the end result is a block cipher "with a strength equal to or better than that of Triple-DES and significantly improved efficiency." With the minimum requirements for the key sizes it is clear that an exhaustive key search will be infeasible for many years. Also with a block size of 128 bits the matching ciphertext attack requires a huge number of about 2^{64} ciphertext blocks to come into play.

The submitters of most of the algorithms claim a very high level of security. An exhaustive search for the key is often claimed to be the best attack, or it is claimed that an attacker would need all 2^{128} possible inputs and outputs to succeed.

However, we think that once a few candidates have been selected by NIST, the increased attention of the worlds cryptanalysts will result in new analysis and in levels of security much lower than claimed by the designers. In particular, we conjecture that (theoretical) key-recovery attacks with complexities in the neighborhood of 2^{100} or less will be found against most of the candidates (provided that they are looked at) in 5 to 10 years and therefore with a security level lower than the best known key-recovery attacks on triple-DES today. Also, a long-time conjecture is that the (theoretical) security level of the final candidate, or the final few candidates in case NIST should decide for several algorithms, will drop to less than 2^{64} in 30 years from now.

8 Conclusion and Open Problems

This paper considers contemporary block ciphers. In the last decade there has been a huge increase in the public knowledge regarding the security of secret-key block ciphers, most notably through the publication of the differential and linear attacks. Although this has enabled us to break many systems faster than by an exhaustive search for the key, the best known attacks on many of these systems are not very practical and require either the encryptions of unrealisticly many chosen or known plaintexts and/or a huge memory and processing time. The open problems in cryptanalysis of block ciphers are easy to spot: Break all unbroken block ciphers! And there is a lot of them.

References

1. K. Aoki and K. Ohta. Differential-linear attack on FEAL. *IEICE Trans. Fundamentals*, E79-A(1):20–27, 1996. 117
2. 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. 118

3. E. Biham. On Matsui's linear cryptanalysis. In A. De Santis, editor, *Advances in Cryptology: EUROCRYPT'94, LNCS 950*, pages 341–355. Springer Verlag, 1995. 116
4. E. Biham and A. Biryukov. An improvement of Davies' attack on DES. In A. De Santis, editor, *Advances in Cryptology: EUROCRYPT'94, LNCS 950*, pages 461–467. Springer Verlag, 1995. 116
5. E. Biham, A. Biryukov, and A. Shamir. "Impossible" cryptanalysis. Presented at the rump session of CRYPTO'98. 114, 115
6. E. Biham and A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*. Springer Verlag, 1993. 112, 113, 115
7. E. Biham and A. Shamir. Differential cryptanalysis of the full 16-round DES. In E.F. Brickell, editor, *Advances in Cryptology: CRYPTO'92, LNCS 740*, pages 487–496. Springer Verlag, 1993. 113
8. M. Blaze, W. Diffie, R.L. Rivest, B. Schneier, T. Shimomura, E. Thompson, and M. Wiener. Minimal key lengths for symmetric ciphers to provide adequate commercial security. Document, January 1996. 120
9. J.B. Borst, L.R. Knudsen, and V. Rijmen. Two attacks on IDEA. In W. Fumy, editor, *Advances in Cryptology: EUROCRYPT'97, LNCS 1233*, pages 1–13. Springer Verlag, 1997. 112, 114, 115
10. P.M. Cohn. *Algebra, Volume 1*. John Wiley & Sons, 1982. 117
11. D. Coppersmith. The real reason for Rivest's phenomenon. In H.C. Williams, editor, *Advances in Cryptology: CRYPTO'85, LNCS 218*, pages 535–536. Springer Verlag, 1986. 118
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. 111
13. J. Daemen, R. Govaerts, and J. Vandewalle. Weak keys for IDEA. In D.R. Stinson, editor, *Advances in Cryptology: CRYPTO'93, LNCS 773*, pages 224–231. Springer Verlag, 1993. 118
14. I.B. Damgård and L.R. Knudsen. The breaking of the AR hash function. In T. Helleseth, editor, *Advances in Cryptology: EUROCRYPT'93, LNCS 773*, pages 286–292. Springer Verlag, 1993. 118
15. I.B. Damgård and L.R. Knudsen. Multiple encryption with minimum key. In E. Dawson and J. Golic, editors, *Cryptography: Policy and Algorithms. International Conference, Brisbane, Queensland, Australia, July 1995, LNCS 1029*, pages 156–164. Springer Verlag, 1995. 110
16. I.B. Damgård and L.R. Knudsen. Two-key triple encryption. *The Journal of Cryptology*, 11(3):209–218, 1998. 110, 119, 121
17. D. Davies and S. Murphy. Pairs and triples of DES S-boxes. *The Journal of Cryptology*, 8(1):20–27, 1995. 116
18. D.W. Davies and W.L. Price. *Security for Computer Networks*. John Wiley & Sons, 1989. 105, 106, 107
19. D.E. Denning. *Cryptography and Data Security*. Addison-Wesley, 1982. 112
20. W. Diffie and M. Hellman. New directions in cryptography. *IEEE Trans. on Information Theory*, IT-22(6):644–654, 1976. 105
21. W. Diffie and M. Hellman. Exhaustive cryptanalysis of the NBS data encryption standard. *Computer*, pages 74–84, 1977. 121
22. C. Harpes, G.G. Kramer, and J.L. Massey. A generalization of linear cryptanalysis and the applicability of Matsui's piling-up lemma. In L. Guillou and J.-J. Quisquater, editors, *Advances in Cryptology - EUROCRYPT'95, LNCS 921*, pages 24–38. Springer Verlag, 1995. 117

23. C. Harpes and J.L. Massey. Partitioning cryptanalysis. In E. Biham, editor, *Fast Software Encryption, Fourth International Workshop, Haifa, Israel, January 1997, LNCS 1267*, pages 13–27. Springer Verlag, 1997. 117
24. M. Hellman. A cryptanalytic time-memory trade-off. *IEEE Trans. on Information Theory*, IT-26(4):401–406, 1980. 111
25. M.E. Hellman and S.K. Langford. Differential–linear cryptanalysis. In Y. Desmedt, editor, *Advances in Cryptology: CRYPTO'94, LNCS 839*, pages 26–39. Springer Verlag, 1994. 117
26. T. Jakobsen. Cryptanalysis of block ciphers with probabilistic non-linear relations of low degree. In H. Krawczyk, editor, *Advances in Cryptology: CRYPTO'98, LNCS 1462*, pages 212–222. Springer Verlag, 1998. 117
27. 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. 114, 117, 119
28. D. Kahn. *The Codebreakers*. MacMillan, 1967. 105, 108
29. B.S. Kaliski and M.J.B. Robshaw. Linear cryptanalysis using multiple approximations. In Y. Desmedt, editor, *Advances in Cryptology: CRYPTO'94, LNCS 839*, pages 26–39. Springer Verlag, 1994. 116
30. 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. 118, 119, 121
31. J. Kilian and P. Rogaway. How to protect DES against exhaustive key search. In Neal Koblitz, editor, *Advances in Cryptology: CRYPTO'96, LNCS 1109*, pages 252–267. Springer Verlag, 1996. 121
32. L.R. Knudsen. Block ciphers - a survey. To appear in the proceedings of the International Course on the State of the Art and Evolution on Computer Security and Industrial Cryptography 1997, to be published in the LNCS Series from Springer Verlag. 113
33. 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. 118
34. L.R. Knudsen. *Block Ciphers – Analysis, Design and Applications*. PhD thesis, Aarhus University, Denmark, 1994. 111, 112, 120
35. L.R. Knudsen. Practically secure Feistel ciphers. In R. Anderson, editor, *Fast Software Encryption - Proc. Cambridge Security Workshop, Cambridge, U.K., LNCS 809*, pages 211–221. Springer Verlag, 1994. 121
36. 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. 118
37. 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. 114, 115
38. 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. 114
39. L.R. Knudsen and T. Berson. Truncated differentials of SAFER. In Gollmann D., editor, *Fast Software Encryption, Third International Workshop, Cambridge, UK, February 1996, LNCS 1039*, pages 15–26. Springer Verlag, 1995. 115

40. L.R. Knudsen and W. Meier. Improved differential attack on RC5. In Neal Koblitz, editor, *Advances in Cryptology - CRYPTO'96, LNCS 1109*, pages 216–228. Springer Verlag, 1996. 114, 118
41. L.R. Knudsen and M.P.J. Robshaw. Non-linear approximations in linear cryptanalysis. In U. Maurer, editor, *Advances in Cryptology: EUROCRYPT'96, LNCS 1070*, pages 224–236. Springer Verlag, 1996. 116
42. X. Lai. On the design and security of block ciphers. In J.L. Massey, editor, *ETH Series in Information Processing*, volume 1. Hartung-Gorre Verlag, Konstanz, 1992. 114
43. 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. 114
44. X. Lai, J.L. Massey, and S. Murphy. Markov ciphers and differential cryptanalysis. In D.W. Davies, editor, *Advances in Cryptology - EUROCRYPT'91, LNCS 547*, pages 17–38. Springer Verlag, 1992. 113, 115
45. J.L. Massey. Cryptography: Fundamentals and applications. Copies of transparencies, Advanced Technology Seminars, 1993. 109, 119
46. J.L. Massey. SAFER K-64: A byte-oriented block-ciphering algorithm. In R. Anderson, editor, *Fast Software Encryption - Proc. Cambridge Security Workshop, Cambridge, U.K., LNCS 809*, pages 1–17. Springer Verlag, 1994. 115
47. J.L. Massey. SAFER K-64: One year later. In B. Preneel, editor, *Fast Software Encryption - Second International Workshop, Leuven, Belgium, LNCS 1008*, pages 212–241. Springer Verlag, 1995. 115
48. 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. 115, 116
49. M. Matsui. The first experimental cryptanalysis of the Data Encryption Standard. In Y.G. Desmedt, editor, *Advances in Cryptology - CRYPTO'94, LNCS 839*, pages 1–11. Springer Verlag, 1994. 112, 116
50. 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. 120
51. 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. 115
52. A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1997. 119
53. 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. 115
54. K. Nyberg and L.R. Knudsen. Provable security against a differential attack. *The Journal of Cryptology*, 8(1):27–38, 1995. 114, 115, 120
55. National Bureau of Standards. Data encryption standard. Federal Information Processing Standard (FIPS), Publication 46, National Bureau of Standards, U.S. Department of Commerce, Washington D.C., January 1977. 106
56. 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. 106
57. National Institute of Standards and Technology. Advanced encryption algorithm (AES) development effort. `http://www.nist.gov/aes`. 106, 122

58. National Institute of Standards and Technology. AES candidate algorithms. Descriptions available from NIST, see `http://www.nist.gov/aes`. 122
59. B. Preneel. *Analysis and Design of Cryptographic Hash Functions*. PhD thesis, Katholieke Universiteit Leuven, January 1993. 118
60. V. Rijmen. *Cryptanalysis and Design of Iterated Block Ciphers*. PhD thesis, Katholieke Universiteit Leuven, October 1997. 112
61. V. Rijmen, B. Preneel, and E. De Win. On weaknesses of non-surjective round functions. *Designs, Codes, and Cryptography*, 12(3):253–266, 1997. 118
62. R. Rivest. The RC5 encryption algorithm. In B. Preneel, editor, *Fast Software Encryption - Second International Workshop, Leuven, Belgium, LNCS 1008*, pages 86–96. Springer Verlag, 1995. 110, 118
63. B. Schneier. Description of a new variable-length key, 64-bit block cipher (Blowfish). In R. Anderson, editor, *Fast Software Encryption - Proc. Cambridge Security Workshop, Cambridge, U.K., LNCS 809*, pages 191–204. Springer Verlag, 1994. 110
64. C.E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28:656–715, 1949. 108, 109, 119
65. M.E. Smid and D.K. Branstad. The Data Encryption Standard: Past and future. In G.J. Simmons, editor, *Contemporary Cryptology - The Science of Information Integrity*, chapter 1, pages 43–64. IEEE Press, 1992. 106
66. W. Tuchman. Hellman presents no shortcut solutions to DES. *IEEE Spectrum*, 16(7):40–41, July 1979. 121
67. S. Vaudenay. An experiment on DES - statistical cryptanalysis. In *Proceedings of the 3rd ACM Conferences on Computer Security, New Delhi, India*, pages 139–147. ACM Press, 1995. 112, 117
68. S. Vaudenay. On the weak keys of Blowfish. In Gollmann D., editor, *Fast Software Encryption, Third International Workshop, Cambridge, UK, February 1996, LNCS 1039*, pages 27–32. Springer Verlag, 1996. 118
69. M.J. Wiener. Efficient DES key search. Technical Report TR-244, School of Computer Science, Carleton University, Ottawa, Canada, May 1994. Presented at the Rump Session of CRYPTO'93. 119
70. M.J. Wiener. Efficient DES key search - an update. *CryptoBytes*, 3(2):6–8, 1998. 119