DES-80*

Carlisle M. Adams
Entrust Technologies
750 Heron Road
Ottawa, Canada, K1V 1A7

1. Abstract

In the Fall of 1996, the Canadian Government issued a request for a study on the feasibility of strengthening the Data Encryption Standard (DES) by increasing the key length to 80 bits. What made this request both interesting and challenging was the overall constraint placed on the project: there was to be no change whatsoever made to the actual encryption / decryption algorithm; rather, changes were to be confined solely to the key scheduling algorithm. Any method may be used to input and process the (maximum) 80-bit primary key, but the result of the process must be sixteen 48-bit round keys suitable for keying the DES rounds in the standard manner.

This paper summarizes the results of the above study, including a new key scheduling algorithm that appears to satisfy all requirements of the DES-80 project. The full report (dated May 2, 1997) is the property of the Canadian Government.

2. Introduction

The Data Encryption Standard (DES) is perhaps the most widely known and widely used cryptographic algorithm in the world today. Since its introduction to the public in the mid-1970s, it has undergone intensive scrutiny by academics, government agencies, industry, and a host of would-be cryptanalysts (both the serious and the hobbyist). Two decades of such focused attention has convinced many that the algorithm itself is basically sound and that the principles used in its design have resulted in a cipher with intrinsic cryptographic strength.

Unfortunately, however, the DES algorithm as originally proposed has a keysize that is too small for some environments. Furthermore, with the rapid advances in computing performance, even on relatively low-priced desktop machines, it is clear that this keysize will become too small for many (or most) environments in the fairly near future. It appears that only two alternatives are possible if security is to be maintained: find a suitable DES replacement algorithm that has cryptographic strength commensurate with the predicted need for the next several years; or modify DES itself in order to increase its strength.

Although the first alternative is being pursued vigorously by many researchers (and has resulted in a number of candidate ciphers), it is recognized that the probability is relatively low that any other cipher will undergo the length and breadth of analysis that DES has so far undergone. Consequently, it may take a considerable length of time before any significant degree of confidence in the security of any given cipher is gained.

The second alternative is the ultimate goal of the DES-80 project. The intention is to leave the algorithm entirely as originally proposed, thereby drawing on the confidence built up over two decades of intense DES-related study. It is hoped that by altering the key schedule alone it may be possible to increase the key length of DES without weakening the cipher in any respect.

The remainder of this paper is organized as follows. Section 3 discusses key schedule design criteria for DES-like ciphers. Section 4 provides a number of candidate proposals for the DES-80 key schedule, and Section 5 evaluates these proposals with respect to the design criteria and the constraints of the DES-80 project. Section 6 describes a new key scheduling algorithm designed specifically for this project and compares it with the other candidates. The paper closes in Section 7 with some concluding remarks.

3. Key Schedule Design Criteria

Over the years a number of researchers have discussed and proposed design criteria for key scheduling algorithms in DES-like ciphers and/or have attempted to derive from the specification of DES the criteria that went

* The work described in this paper was funded by the Communications Security Establishment of the Canadian Government.

into its design. This section summarizes that work and lists the criteria that appear to be important in key schedule design.

3.1 Bit Effectiveness

This criterion strives to avoid ineffective key bits (for any arbitrarily-chosen primary key) by ensuring that each bit is used as an input to each s-box somewhere in the cipher (e.g., since DES uses 8 s-boxes, then each key bit must be input to each of the 8 s-boxes by the time the final round has been computed) [6, 14]. This general objective is refined slightly in [1], where it is recommended that all key bits be used by the time that half the rounds have been computed, and reused, possibly in a different order, in the lower half of the network (thus ensuring good key avalanche for both encryption and decryption). The objective is slightly further refined in [7], where it is suggested that each key bit should be presented to each s-box "as quickly as possible".

3.2 Bit Diffusion

If key bits are reused through the rounds (if the key is anything other than 768 truly independent bits in DES, for example), then a given bit should have a different "role" with each use. This concept has been expressed as follows: no key bit is used as input to the same s-box on successive rounds [6, 14]. Note that this criterion can therefore satisfy the "bit effectiveness" criterion (Section 3.1) for any cipher that has more rounds than s-boxes (such as DES).

In some ways, this criterion may be seen as an extreme expression of the "rotating keys" criterion (Section 3.11).

3.3 Key / Ciphertext Completeness

Kam and Davida [9] defined the property of plaintext/ciphertext "completeness" for substitution-permutation network cryptosystems having an n-bit blocksize and an m-bit keysize as follows: for every possible key value, every output bit c_i of the SP network depends on all input bits $p_1, p_2, ..., p_n$, and not just a proper subset of the input bits. It is not difficult to construct a similar definition for key/ciphertext completeness (which should also be a property of any good cipher, as noted in [1]): for every possible plaintext value, every output bit c_i of the SP network depends on all key bits $k_1, k_2, ..., k_m$, and not just a proper subset of the key bits.

A number of researchers have reiterated this general requirement by stating that dependency of every ciphertext bit on all key bits should increase rapidly through the rounds; with respect to DES, analysis has shown that every ciphertext bit is dependent on all key bits after five rounds [7, 14, 15].

3.4 Key / Ciphertext SAC, BIC

Webster and Tavares [19, 20] defined the property of Strict Avalanche Criterion (SAC) as follows: output bit j should change with probability ½ when any single input bit i is inverted, for all i, j (where the probability is computed over the set of all pairs of input vectors that differ only in bit i). The property of (output) Bit Independence Criterion (BIC) was defined similarly: output bits j and k should change independently when any single input bit i is inverted, for all i, j, k (where the independence is computed over the set of all pairs of input vectors that differ only in bit i).

"Higher orders" of SAC and BIC involve multiple-bit changes to the input and, for the case of BIC, also involve combinations of multiple output bits (see, for example, the definitions and notation given in [16]).

Applying these concepts to the topic of key scheduling leads to the criterion that the key schedule should satisfy (at least to a close approximation) SAC, BIC, higher-order SAC, and higher-order BIC. As well (like the key/ciphertext completeness criterion of Section 3.3), this dependency should increase rapidly through the rounds.

3.5 Pseudo-Independence of "Intra" Round Key Bits

Knudsen [11] (see also [5]) has proposed the following property of a "strong" key schedule: given any s bits of the set of round keys, derived from an unknown primary key (where s is less than the total amount of round key material), it is *hard* to find any of the remaining round key bits from the s known bits. The term *hard* may be replaced by a more precise definition depending on the application, but it is noted as a practical limit that it cannot be harder than performing the key schedule for all possible primary keys.

3.6 Pseudo-Independence of "Inter" Round Key Bits

Knudsen [11] (see also [5]) has proposed the following property of a "strong" key schedule: given some relation between two primary keys, it is *difficult* to predict the relations between any of the round keys derived from these two primary keys. The term *difficult* may be replaced by a more precise definition depending on the

application, but it is noted as a practical limit that it cannot be more difficult than performing the key schedule for the two primary keys. Furthermore, if the primary keys are unknown (i.e., their relation is all that is known), then *difficult* cannot be more difficult than performing the key schedule for all possible pairs of primary keys.

3.7 Absence of Weak and Semi-Weak Keys

Weak and semi-weak keys for DES have been defined and examined in a number of places (see, for example, [8, 15, 17, 18]). Weak keys have the property that double encryption returns the original plaintext (i.e., $E_k(E_k(p)) = p$, where k is a weak key and p is any plaintext). Semi-weak keys also have the property that double encryption returns the original plaintext, but in this case each encryption uses a different key (i.e., $E_{k2}(E_{k1}(p)) = p$ (where $k1$ and $k2$ are a semi-weak pair and p is any plaintext). DES has 4 weak keys and 12 semi-weak keys (6 semi-weak pairs).

It is clear that such a small number of weak and semi-weak keys implies that these are extremely unlikely to be chosen by chance in any given environment, so many implementations do not check for these when randomly generating DES keys. However, it should be noted that while weak keys create few problems for enciphering (since they are unlikely to occur), they can be a problem for hash functions based on DES because the key input may be chosen by the attacker in attempts to find collisions. Thus, it is concluded that an important design criterion of a key schedule is that it produces no weak or semi-weak keys [11].

3.8 No Easily-Found Fixed Points

A fixed point of a key k is a plaintext vector x such that $E_k(x) = x$, and an anti-fixed point of a key k is a plaintext vector x such that $E_k(x)$ is the complement of x (see [17, 18], for example). It is known for DES that each of the weak keys has 2^{32} easily-found fixed points and each of four of the semi-weak keys has 2^{32} easily-found anti-fixed points (where "easily-found" means that a level of effort of roughly 2^{32} operations, rather than 2^{64} operations, is required).

Since easily-found fixed points and anti-fixed points are a consequence of the details of the key scheduling algorithm, a reasonable design goal for a new key schedule is that it should have no easily-found fixed or anti-fixed points for any key. From all evidence available thus far in the open literature, fixed and anti-fixed points have only been easily found in DES-like ciphers for weak and semi-weak keys. It is therefore conjectured that a key schedule will satisfy this criterion if it can provably avoid producing weak and semi-weak keys (the criterion of Section 3.7) [2].

3.9 Absence of Quasi-Weak Keys

Knudsen [13] has shown that for a large number of pairs of keys in DES (so-called "quasi-weak" keys) there is a simple relation between the encryption functions induced by these keys. This relation is a result of the fact that these quasi-weak key pairs have a significant number of round keys in common (e.g., 12 or 13). Although these keys appear not to be a problem for encryption / decryption in typical applications, it is noted that some concern may exist for hash functions based on DES (in which the keys are fixed or can be chosen as part of the hash message).

A desirable criterion for the DES-80 key schedule, therefore, is that it should avoid quasi-weak keys.

3.10 Absence of Related Keys

Biham [3] has shown the importance of avoiding key schedules that produce obvious relationships between round keys by describing chosen-plaintext attacks that are of lower complexity than exhaustive search, and low-complexity chosen-key attacks, both of which are independent of the number of rounds of the cipher and of the details of the round function.

He notes that DES is not vulnerable to the attacks described because of the irregularity of the shift pattern in its key schedule. An important criterion for the DES-80 key schedule, therefore, is that it should avoid related keys (although it need not necessarily do this through the use of an irregular shift pattern).

3.11 Need for Rotating Keys

Grossman and Tuckerman [8] have shown that a Feistel-like cipher that does not use a "rotating key" can be broken. That is, if the same key bits are used in the same way for every round, then the cipher will succumb to a chosen plaintext attack, regardless of the number of rounds. This leads to the criterion that the key schedule must have rotating keys (i.e., round keys that are not identical over successive rounds).

Note that the criteria of "bit effectiveness" and "bit diffusion" (Sections 3.1 and 3.2, resp.) strengthen and make more explicit the intention of this criterion.

3.12 Absence of Complementation Property

It is well known that DES has the "complementation" property: $E_k(p) = \overline{E_{\bar{k}}(\bar{p})}$ (see [15], for example). Although this does little harm in terms of security (reducing the work factor of exhaustive key search from 2^{56} to 2^{55} encryptions), it seems highly unlikely that possession of this property was an explicit design goal for DES; it is much more probable that this was simply an unfortunate side effect of the way the key scheduling is done and the way the round keys are used within the cipher.

Given, however, the opportunity to completely re-design the key schedule for DES, it seems that there is little reason to maintain this long-standing side effect. An explicit design goal of the DES-80 project, therefore, is that the revised cipher should not possess the complementation property.

3.13 Decryption Considerations

The current DES key schedule uses rotations as one aspect of its computation of successive round keys. The total number of bits rotated through the schedule is 56, which implies that $K(0) = K(16)$; this enables the decryption operation to use right shifts in reverse order (as opposed to left shifts in "forward order" – the order specified in DES) [6]. The current key schedule thus allows the first round key for decryption to be computed as quickly as the first round key for encryption, which may lead to a slightly decreased set-up time before decryption can begin in some specialized implementations.

In typical implementations, however, the full set of round keys is computed before encryption or decryption is begun, so a more practical criterion is that the computation of the set of round keys for decryption should not be significantly more time consuming than the computation of the set of round keys for encryption. (But note that this more relaxed criterion is readily fulfilled by any key schedule, since the round keys can always be computed in "forward" order and then *used* in reverse order for decryption; that is, they need never actually be *computed* in reverse order.)

3.14 Round Key Set-Up Time

The time required to establish the complete set of round keys should not be significantly different from that required with the current DES key scheduling algorithm (for encryption as well as for decryption). This is particularly important for environments wherein very short messages need to be encrypted/decrypted or keys need to be changed relatively frequently.

3.15 Implementation Simplicity

Although not strictly important in terms of security, for practical considerations it is desirable that the key schedule be relatively simple to implement. This can lead to implementations that are completed more quickly and that are more likely to be correct the first time, which in turn leads to cost savings and other associated benefits.

3.16 Summary

With respect to the criteria listed above, it can be said that the current DES key schedule fails Sections 3.5, 3.6, 3.7, 3.8, 3.9, and 3.12 (to varying degrees), and essentially satisfies the remaining Sections. An explicit design intent for the DES-80 key schedule, of course, is to satisfy as many of the above criteria as possible.

From the descriptions given above it is not difficult to see that Sections 3.1 and 3.11 are subsumed by 3.2, 3.3 is subsumed by 3.4, 3.8 appears to be satisfied by 3.7, and 3.13 appears to be of little practical value to many implementations. Thus, the following criteria are judged to be of most relevance and importance to the DES-80 project:

Section 3.2: diffusion of key bits;

Section 3.4: key/ciphertext SAC, BIC (and higher orders);

Section 3.5: pseudo-independence of intra-round keys;

Section 3.6: pseudo-independence of inter-round keys;

Section 3.7: absence of weak and semi-weak keys;

Section 3.9: absence of quasi-weak keys;

Section 3.10: absence of related keys;

Section 3.12: absence of complementation property;

Section 3.14: short round key set-up time; and

Section 3.15: implementation simplicity.

These, then, are the criteria that are used for evaluation of the various key schedule proposals in Section 5.

4. Candidates from the Open Literature

In this section some of the most significant/important key scheduling proposals are described in some detail (the full report contains additional proposals that are not included in this summary due to lack of space). The proposals are examined in the following section with respect to both the design criteria distilled from the literature and the constraints imposed by the DES-80 project itself.

4.1 Knudsen (1994)

Knudsen [11] examines two main areas of Feistel cipher design: "strong" key scheduling; and "nonlinear and differentially uniform" round function construction. In the discussion on key scheduling, two proposals are given (the first keeps the key length at 56 bits but increases resistance to linear, differential, and exhaustive search cryptanalysis; the second doubles the key length (to 112 bits, thus rendering exhaustive search computationally infeasible) and has the same increase in resistance to linear and differential attacks as the first proposal). Knudsen recommends the second proposal in recognition of the fact that DES requires a larger key for many environments.

Proposal

The proposal is as follows:

$$RK_j = 48MSB(DES_{K1}(DES_{K2}^{-1}(DES_{K1}(IV \oplus j)))).$$

Thus, two-key triple-DES is used to compute the round keys, leading to a 112-bit version of DES.

4.2 Biham, Biryukov (1995)

The paper by Biham and Biryukov [4] explores and analyzes four main approaches for extending the key length of DES: key-dependent s-box *transformations*; key-dependent s-box *reorderings*; key-dependent s-box "*choices*"; and key-dependent s-box "*contents*". The first creates new s-boxes through the transformation of existing s-boxes (by XORing key material before and after the s-box look-up). The second uses key bits to choose an order of the s-boxes in the round function. The third uses key bits to select new s-boxes for the round function. Finally, the fourth uses key bits to generate s-boxes with random (or pseudorandom) content.

Proposal

Biham and Biryukov propose the following modifications to DES. Their key is of the form (K_a, K_b, K_c, K_d), where each component is specified as follows.

> K_a consists of 16 bits, expanded (as specified in [4]) to 48 bits; these are XORed to the input of the s-boxes in all the rounds.
>
> K_b consists of 32 bits; these are XORed to the output of the s-boxes in all the rounds.
>
> K_c consists of 8 bits if DES s-boxes are used and consists of 15 bits if s^{3}DES (see [11]) s-boxes are used; these choose an order in which the s-boxes are loaded (where each combination of bits corresponds to one of the strong orders of the s-boxes).
>
> K_d consists of 56 bits; these are loaded into the original DES key scheduling algorithm.

Thus, the key consists of 56+16+32+8 = 112 bits in the case of DES s-boxes and 119 bits in the case of s^{3}DES s-boxes (note that the authors recommend the s^{3}DES s-boxes – with a reversed order of S1 and S2 – since these appear to be more resistant to both linear and differential cryptanalysis than the original DES s-boxes). The keys K_a, K_b, and K_c are used to construct a set of s-boxes from the starting set; this new set is then loaded into the DES algorithm along with K_d. Key scheduling is done with K_d as specified in the original DES and encryption or decryption proceeds as in DES (but with the new s-boxes).

4.3 Kilian, Rogaway (1996)

Kilian and Rogaway [10] show that the DESX construction is effective in protecting DES against exhaustive key search. Specifically, let κ be the key length for a block cipher and let n be its block length. An *ideal* block

cipher with these parameters is modeled as a *random* map $F: \{0,1\}^{\kappa} \times \{0,1\}^{n} \rightarrow \{0,1\}^{n}$ subject to the constraint that for every key $k \in \{0,1\}^{\kappa}$, $F_k(\cdot)$ is a permutation on $\{0,1\}^{n}$. This ideal cipher F is then extended to FX, where $FX: \{0,1\}^{\kappa+2n} \times \{0,1\}^{n} \rightarrow \{0,1\}^{n}$, by setting

$$FX_K(x) = FX_{k\|k_1\|k_2}(x) = k_2 \oplus F_k(k_1 \oplus x).$$

It is shown that using *only* exhaustive search (i.e., treating FX as a "black box" and having access only to plaintext/ciphertext pairs without knowing anything whatsoever about the internal structure of the cipher), the effective key length is increased from κ to $\kappa+n-1-\lg m$ bits, where m bounds the number of $\langle x, FX_K(x)\rangle$ pairs that the adversary can obtain.

Proposal

Although it is recognized that DES is not an *ideal* cipher (i.e., a family of *random* permutations), it is also recognized that it does *appear* to be an ideal cipher with respect to pure black box statistical analysis (that is, analysis that does not exploit the internal structure of the cipher) except for its complementation property. Thus, the effective key length of DES is 55 bits and so the effective key length of DESX is 55+64-1-lg$m$ bits. If an adversary can obtain up to $m=2^{30}$ blocks of enciphered data (a reasonable upper bound, since greater amounts will quickly lead to other attacks arising from the birthday paradox), then DESX has an effective key length of 88 bits.

Kilian and Rogaway therefore recommend the use of DESX wherever DES is currently used since it is efficient, DES-compatible, patent unencumbered, and at least 88 bits strong against exhaustive key search. Furthermore, they note that DESX retains its strength even if $k_1=k_2$ (i.e., if DESX is defined as $k_1 \oplus \mathrm{DES}_k(k_1 \oplus x)$), so that the primary key can be 56+64=120 bits rather than 184 bits. Finally, although it is freely admitted that DESX was never intended to defend against differential or linear cryptanalysis (or indeed against any attack that exploits the internal structure of DES), the authors note that their proofs still hold when "$\oplus$" is replaced by a variety of other binary operations and that, in particular, addition modulo 2^{32} appears to offer some resistance to differential and linear attacks.

4.4 Blumenthal, Bellovin (1996)

Blumenthal and Bellovin [5] advocate the generation of round keys for any symmetric block cipher – and in particular for DES – in such a way that finding relations between any round key bits (over all rounds) is intractable for the foreseeable future. The proposed key schedule has a considerable set-up time (equivalent to 43 encryptions for DES), but the authors feel that this can be helpful in foiling brute force attacks in cases where the primary key is relatively small.

Proposal

The concept of "n-folding" is defined in order to take a variable-length input block and produce a fixed-length output block such that each input bit contributes approximately equally in determining the value of each output bit.

The n-folding operation is described in [5] as follows. To n-fold a number X, replicate the input value to a length that is the least common multiple of n and the length of X. Before each repetition, the input is rotated to the right by 13 bit positions. The successive n-bit chunks are added together using 1's complement addition (that is, addition with end-around carry) to yield an n-bit result denoted $\langle X \rangle_n$.

Let the primary key K be of any desired length. Let $E_k(P)$ represent encryption of plaintext P using DES in ECB mode with key k (and let $D_k(P)$ represent the corresponding decryption); let $E_{k;IV}(P)$ represent encryption of plaintext P using DES in CFB-64 mode with key k and initialization vector IV; and let $E_{\{h\};IV}(P)$ represent encryption of plaintext P using 3-key triple-DES in CFB-64 mode with key $\{h\} = h_1,h_2,h_3$ and initialization vector IV. Compute

$$A = \langle K \rangle_{64}$$
$$IV = E_{E_A(A)}(A)$$
$$k = D_{D_A(A)}(A)$$
$$\{h\} = E_{k;IV}(\langle K \rangle_{168})$$
$$R = E_{\{h\};IV}(\langle K \rangle_{768})$$

The 768-bit value R is then used as the bits of the full set of round keys. This computation requires 2 DES encryptions in ECB mode (for IV), 2 decryptions in ECB mode (for k), 3 encryptions in CFB mode (for $\{h\}$), and 12 triple-DES encryptions in CFB mode (for R), for a total of 43 encryptions (plus some extra time for the 64-, 168-, and 768-folding operations, along with 7 DES key scheduling operations (1 for A, 1 for $E_A(A)$, 1 for $D_A(A)$, 1 for k, and 3 for $\{h\}$).

5. Evaluations

This section provides an analysis of two of the specific proposals given previously (analysis of the others is presented in the full report). The analysis for each proposal includes an evaluation with respect to the key schedule design criteria from Section 3, where a specific proposal is said to **satisfy** a criterion if it performs as well as original DES or better, and is said to **fail** the criterion otherwise. An evaluation is also provided with respect to the constraints imposed by the DES-80 project and each proposal is assessed according to whether it can be said to **meet** or **miss** each constraint. The important constraints under consideration for the DES-80 project are as follows:

localization: changes are to be made to the key scheduling algorithm only (no other parts of the DES algorithm are to be modified in any way);

strength: the resulting DES-80 algorithm is to have equivalent or increased resistance to known attacks (compared with the original DES algorithm);

bounded entropy: the key of the DES-80 algorithm may be fixed or variable in size (and, in fact, a variable size may be preferable for a number of environments), but it must have a maximum length of 80 bits.

5.1 Knudsen (1994)

General Comments

The concrete proposal requires 16 triple-DES encryptions, in terms of set- up time, to generate the round keys.

In order to achieve interoperability, either IV has to be standardized (i.e., fixed in a publicly-accessible way) or it needs to carried as a (not necessarily secret) part of the primary key (thus making the key 64 bits longer than is explicitly necessary).

Analysis with respect to Design Criteria

Diffusion of Key Bits: This proposal **satisfies** this criterion because key bits are not used directly in the rounds, but rather are used to (strongly) pseudorandomly generate round keys.

Key/Ciphertext SAC, BIC: This proposal **satisfies** this criterion because key bits are used to (strongly) pseudorandomly generate round keys (thus, a change in any key bit(s) will lead to large, unpredictable changes in every round key).

Intra-Round Keys: This proposal **satisfies** this criterion because key bits are used to (strongly) pseudorandomly generate round keys (thus, it appears to require breaking triple-DES to find unknown round key bits).

Inter-Round Keys: This proposal **satisfies** this criterion because key bits are used to (strongly) pseudorandomly generate round keys (thus, it appears to require breaking triple-DES to find relationships between round key bits from different primary keys).

Weak, Semi-Weak Keys: This proposal **satisfies** this criterion since it is extremely unlikely that the triple-DES computation will generate palindromic or anti-palindromic sets of keys.

Quasi-Weak Keys: This proposal **satisfies** this criterion since it is extremely unlikely that the triple-DES computation will generate two sets of round keys that have a significant number of round keys in common.

Related Keys: This proposal **satisfies** this criterion since it is extremely unlikely that the triple-DES computation will generate two sets of round keys with obvious relations between corresponding pairs of round keys in the sets.

Complementation Property: This proposal **satisfies** this criterion since the triple-DES computation ensures that complementing the primary key does not simply complement each round key.

Set-Up Time: This proposal **fails** this criterion since it requires the equivalent of 48 encryptions and 3 key schedule operations to generate the round keys.

Implementation Simplicity: This proposal **satisfies** this criterion since only doing triple-DES requires a fairly small implementation effort beyond what would already be required for DES.

Analysis with respect to DES-80 Constraints

Localization: This proposal **meets** this constraint because changes are confined to the key scheduling algorithm alone.

Strength: This proposal **meets** this constraint because resistance to linear, differential, and exhaustive key cryptanalysis is substantially improved and no other cryptanalytic attack is known to be made any easier.

Bounded Entropy: This proposal **misses** this constraint because it is not clear how to turn this 112-bit key schedule into an 80-bit key schedule (primarily because it is not clear that a DES key with $b<56$ randomly-chosen bits, padded out in a known, deterministic way to 56 bits, is *necessarily* "b" bits strong). Even if this can be resolved, however, it would still not be clear how to *restrict* a primary key to 80 bits (i.e., other controls – outside the cipher itself – would need to be put in place to ensure that longer keys cannot be used).

5.2 Biham, Biryukov (1995)

General Comments

It would be possible to produce an 80-bit proposal by eliminating K_b and using the DES s-boxes (no other arrangement seems to readily give 80 bits), but with the two complementation properties mentioned, this is effectively a 78-bit DES.

It is possible to achieve backward compatibility with this proposal (if the DES s-boxes are used and if an all-zero K_c specifies the "standard" order of these s-boxes), which may be a useful feature in some environments.

Analysis with respect to Design Criteria

Diffusion of Key Bits: This proposal **satisfies** this criterion (as does DES) for all key bits in K_d. Since the bits in K_a, K_b, and K_c essentially choose s-boxes and do not define round keys, they are not relevant to this criterion (i.e., they are not re-used through the rounds in the typical sense of round keys).

Key/Ciphertext SAC, BIC: This proposal clearly **satisfies** this criterion as well as DES for all key bits in K_d. This seems, intuitively, to also be true for all key bits in K_c, since rearranging the order of the s-boxes should lead to an unpredictable change in the ciphertext. The criterion should also be satisfied for all bits in K_a and K_b (since these swap rows or complement columns in s-boxes), but this seems less obviously true and may require some further analysis.

Intra-Round Keys: This proposal **fails** this criterion (as does DES) for all key bits in K_d. Since the bits in K_a, K_b, and K_c essentially choose s-boxes and do not define round keys, they are not relevant to this criterion.

Inter-Round Keys: This proposal **fails** this criterion (as does DES) for all key bits in K_d. Since the bits in K_a, K_b, and K_c essentially choose s-boxes and do not define round keys, they are not relevant to this criterion.

Weak, Semi-Weak Keys: This proposal **fails** this criterion because it has $2^{56}x2^4=2^{60}$ weak and semi-weak keys instead of 2^4 (since it does not matter what values K_a, K_b, and K_c have). Note, however, that it is still not worth screening for them, since there is no higher chance of generating them at random than there is with original DES.

Quasi-Weak Keys: This proposal **fails** this criterion (as does DES) for the key bits in K_d. Further analysis is required to determine whether or not modifying the s-boxes (through K_a, K_b, and K_c) suffices to prevent the possibility of quasi-weak keys, but this seems doubtful.

Related Keys: This proposal **satisfies** this criterion (as does DES) for the key bits in K_d. Since the bits in K_a, K_b, and K_c essentially choose s-boxes and do not define round keys, they are not relevant to this criterion.

Complementation Property: This proposal **fails** this criterion (as does DES) since modifying the s-boxes cannot prevent the complementation property and the remainder of the algorithm is identical to DES.

Set-Up Time: This proposal **satisfies** this criterion conceptually; in practice the criterion is only satisfied if the s-box contents can be manipulated and installed into the round function quickly and easily.

Implementation Simplicity: This proposal **satisfies** this criterion conceptually; in practice the criterion is only satisfied if the s-box contents can readily be manipulated (through K_a, K_b, and K_c) by the implementation.

Analysis with respect to DES-80 Constraints

Localization: This proposal **misses** this constraint because it is completely predicated on the notion that s-boxes can be loaded into the DES implementation; new s-boxes demand an implementation change that is beyond the confines of the key scheduling algorithm. It is worth noting, however, that DES chips do exist in the marketplace that allow loading of s-boxes [4] (although [4] does not mention how time consuming this process might be) and that software implementations may be written in such a way as to dynamically link new s-boxes at run-time, if this is desired.

Strength: This proposal **meets** this constraint because resistance to linear, differential, and exhaustive key cryptanalysis is substantially improved and no other cryptanalytic attack is known to be made any easier.

Bounded Entropy: This proposal **meets** this constraint if K_b is eliminated and if the original DES s-boxes are used. The key size may easily be varied from 56 to 80 bits, inclusive (sizes below 56 bits are possible if and only if full-entropy sizes below 56 bits are possible for original DES; that is, if DES keys with $b<56$ randomly-chosen bits, padded out in a known, deterministic way to 56 bits, are *necessarily* "b" bits strong).

6. A New Key Schedule

This section introduces a new key scheduling algorithm that makes extensive use of the DES round function to pseudorandomly generate a set of round keys from a primary key. This proposal adopts the key scheduling model embraced by [5] and [11], but begins with a primary key of maximum length 80 bits and requires less than two DES encryptions to produce the full set of round keys. The pseudo-independence of the round key bits relies on the statistical properties of the DES round function (which are generally acknowledged to be good). It is conjectured that given certain sets of 4 round keys k_i ... k_{i+4}, it is *hard* to compute the subsequent round keys k_{i+5} ... k_{16}, where "hard" means that a work factor of roughly 2^{48} operations is required (note, however, that if even a *single* round key can be determined in DES, the original cipher is trivially broken, so sets of four round keys are likely to be very difficult to find). Furthermore, it is conjectured that computing previous round keys from one or more known round keys requires roughly 2^{80} operations (i.e., a work factor equivalent to simply doing an exhaustive search on the primary key).

This proposal was motivated by the observation that the DES round function takes exactly 80 bits of input (32 bits of data and 48 bits of key); thus, an 80-bit primary key may be used directly as input to the round function. It is postulated that applying this function iteratively (with an appropriate updating procedure) will produce a pseudorandom stream of bits suitable for constructing the full set of round keys.

Proposal

The following notation is employed:

K = the primary 80-bit key

K_H = the high-order (most significant, leftmost) 32 bits of K

K_L = the low-order (least significant, rightmost) 48 bits of K

f_i = a 32-bit temporary variable

$f_i^{(j)}$ = the j^{th} byte of the variable f_i (where $f_i^{(1)}$ is the least significant byte)

k_i = a DES round key ($i = 1...16$)

"$a <<< b$" = the variable a circularly left-shifted (rotated) by b bits

"$a \| b$" = the variable a concatenated with the variable b

$F(d,r)$ = the DES round function with 32-bit data d and 48-bit round key r

Constructing the set of round keys is a two-step process. The steps are specified in pseudocode as follows.

```
for (i = 1; i ≤ 24; i++)
  {
  f_i = F(K_H,K_L)
  K = (K <<< (i + 8))
  K_H = (K_H + f_i) mod 2^32
  }

for (i = 0; i ≤ 3; i++)
  for (j = 1; j ≤ 4; j++)
    k_4i+j = f_6i+1^(j) || f_6i+2^(j) || f_6i+3^(j) || f_6i+4^(j) || f_6i+5^(j) || f_6i+6^(j)
```

This algorithm computes the DES round function 24 times, modifying its input at each iteration. The 24 32-bit outputs from Step 1 are used in Step 2 to construct the 16 48-bit round keys.

Note that the iteration-dependent rotation in Step 1 (that is, rotation by $(i + 8)$ bits at the i^{th} iteration) helps to defend against a related-key attack on this key schedule. Furthermore, using addition modulo 2^{32} as the

modification operation (rather than XOR) ensures that primary keys that are complements of each other do not lead to identical sets of round keys.

It is not difficult to find two primary keys K and H that produce the same f_1 in Step 1. Furthermore, it is conceivable that K and H can be found that produce both the same f_1 and the same f_2. Therefore, from Step 2, the first byte (or possibly the first two bytes) in the first four corresponding round keys can be made to be identical (that is, $k_1^{(1)} = h_1^{(1)}$, $k_2^{(1)} = h_2^{(1)}$, $k_3^{(1)} = h_3^{(1)}$, $k_4^{(1)} = h_4^{(1)}$, and possibly $k_1^{(2)} = h_1^{(2)}$, $k_2^{(2)} = h_2^{(2)}$, $k_3^{(2)} = h_3^{(2)}$, $k_4^{(2)} = h_4^{(2)}$). Given, however, that the remaining bytes of these four corresponding round keys will be different with very high probability and that all remaining round keys will be statistically independent to virtually any degree of analysis, the above situation does not describe a related-key attack and presents no known security risk.

The key scheduling model embraced by [5] and [11] is also used here because it is recognized that differential cryptanalysis of DES with independent round keys requires 2^{60} chosen plaintexts and finds the 768-bit key in time equivalent to about 2^{61} encryptions (not significantly better than a dictionary attack requiring 2^{64} chosen plaintexts). If an improvement is at all possible against independent round keys, it is conjectured [11] that it would nevertheless require more than the 2^{47} chosen plaintexts used to attack DES with dependent round keys (i.e., with the original key schedule). Furthermore, an estimate for linear cryptanalysis of DES with independent round keys is not known but is conjectured to be high. One approach would be to recover the full round key of the final round and then successively "peel off" rounds until the 768-bit key is recovered, but a linear attack on the final round key will require many linear expressions (including expressions with a probability that requires many known plaintexts) in order to uniquely determine the key. These points argue in favour of independent round keys. However, since 768-bit keys are impractical for most environments, (strongly) pseudo-independent round keys are chosen as an attractive and practical alternative.

General Comments

It is likely that the round keys will look like 768 random bits to any degree of analysis, so the goal of increased resistance to linear, differential, and exhaustive search cryptanalysis with no change to the underlying algorithm appears to be achieved.

This proposal has a set-up time that is comparable to original DES, so it may be amenable to environments in which frequent re-keying is common.

Analysis with respect to Design Criteria

Diffusion of Key Bits: This proposal **satisfies** this criterion because key bits are not used directly in the rounds, but rather are used to pseudorandomly generate round keys.

Key/Ciphertext SAC, BIC: This proposal **satisfies** this criterion because key bits are used to pseudorandomly generate round keys (thus, a change in any key bit(s) will lead to large, unpredictable changes in every round key).

Intra-Round Keys: This proposal **satisfies** this criterion because key bits are used to pseudorandomly generate round keys (thus, it appears to require breaking an iterated DES round function with 80 bits of unknown input to find unknown round key bits).

Inter-Round Keys: This proposal **satisfies** this criterion because key bits are used to pseudorandomly generate round keys (thus, it appears to require breaking an iterated DES round function with 80 bits of unknown input to find relationships between significant numbers of round key bits from different primary keys [see Section 0]).

Weak, Semi-Weak Keys: This proposal **satisfies** this criterion since it is highly unlikely that the iterated DES round function will generate palindromic or anti-palindromic sets of keys.

Quasi-Weak Keys: This proposal **satisfies** this criterion since it is highly unlikely that the iterated DES round function will generate two sets of round keys that have a significant number of round keys in common.

Related Keys: This proposal **satisfies** this criterion since it is highly unlikely that the iterated DES round function will generate two sets of round keys with obvious relations between most or all corresponding pairs of round keys in the sets.

Complementation Property: This proposal **satisfies** this criterion since the iterated DES round function ensures that complementing the primary key does not simply complement each round key.

Set-Up Time: This proposal **satisfies** this criterion since it requires less than 2 encryption operations to establish the set of round keys.

Implementation Simplicity: This proposal **satisfies** this criterion because the iterated DES round function requires little implementation effort beyond what would already be required for DES.

Analysis with respect to DES-80 Constraints

Localization: This proposal **meets** this constraint because changes are confined to the key scheduling algorithm alone.

Strength: This proposal **meets** this constraint because resistance to linear, differential, and exhaustive key cryptanalysis is substantially improved and no other cryptanalytic attack is known to be made any easier.

Bounded Entropy: This proposal **meets** this constraint because the length of the primary key, although variable, is restricted to a maximum of 80 bits.

7. Conclusions

The study summarized in this paper has reviewed and analyzed an extensive list of key scheduling proposals for the Data Encryption Standard. The final recommendation of the study is the new key scheduling proposal presented in Section 6. This proposal appears to meet all requirements and objectives of the DES-80 project. Alternative recommendations, both those that almost meet the DES-80 requirements and those that fall outside the scope of the DES-80 project, are also presented in the full report.

8. References

[1] C. Adams, "A Formal and Practical Design Procedure for Substitution-Permutation Network Cryptosystems", Ph.D. Thesis, Queen's University, Kingston, Canada, 1990.

[3] C. Adams, "Constructing Symmetric Ciphers Using the CAST Design Procedure", *Designs, Codes, and Cryptography* (to appear) [see also the associated "CAST Design Procedure Addendum", located at http://www.entrust.com/library.htm].

[3] E. Biham, "New Types of Cryptanalytic Attacks Using Related Keys", *Journal of Cryptology*, vol. 7, no. 4, Autumn 1994, pp.229-246.

[4] E. Biham and A. Biryukov, "How to Strengthen DES using Existing Hardware", in *Advances in Cryptology: Proceedings of Asiacrypt '94*, Springer-Verlag, 1995, pp.398-412.

[5] U. Blumenthal and S. Bellovin, "A Better Key Schedule for DES-Like Ciphers", in *Advances in Cryptology: Proceedings of Pragocrypt '96*, 1996 (to appear).

[6] L. Brown, "A Proposed Design for an Extended DES", in *Computer Security in the Age of Information: Proceedings of the Fifth International Conference*, North-Holland, 1989, pp.9-22.

[7] L. Brown and J. Seberry, "Key Scheduling in DES Type Cryptosystems", in *Advances in Cryptology: Proceedings of Auscrypt '90*, Springer-Verlag, 1990, pp.221-8.

[8] D. Davies, "Some Regular Properties of the 'Data Encryption Standard' Algorithm", in *Advances in Cryptology: Proceedings of Crypto '82*, Plenum, 1983, pp.89-96.

[9] J. Kam and G. Davida, "Structured Design of Substitution-Permutation Encryption Networks", *IEEE Transactions on Computers*, vol. C-28, 1979, pp.747-753.

[10] J. Kilian and P. Rogaway, "How To Protect DES Against Exhaustive Key Search", in *Advances in Cryptology: Proceedings of Crypto '96*, Springer-Verlag, 1996, pp.252-267.

[11] K. Kim, S. Park, and S. Lee, "Reconstruction of s^2DES S-boxes and their Immunity to Differential Cryptanalysis", *Proceedings of the 1993 Korea-Japan Joint Workshop on Information Security and Cryptology (JW-ISC '93)*, Seoul, Korea, October 24-26, 1993.

[12] L. Knudsen, "Practically Secure Feistel Ciphers", in *Proceedings of the Cambridge Security Workshop on Fast Software Encryption*, Springer-Verlag, 1994, pp.211-221.

[13] L. Knudsen, "New Potentially 'Weak' Keys for DES and LOKI", in *Advances in Cryptology: Proceedings of Eurocrypt '94*, Springer-Verlag, 1995, pp.419-424.

[14] K. Kusuda and T. Matsumoto, "A Strength Evaluation of the Data Encryption Standard", *ISO/TC68 (International Organization for Standardization, Technical Committee TC68) Technical Report*, October, 1996.

[15] C. Meyer and S. Matyas, *Cryptography: A New Dimension in Data Security*, John Wiley & Sons, 1982.

[16] S. Mister and C. Adams, "Practical S-Box Design", in *Workshop Record of the Workshop on Selected Areas of Cryptography (SAC '96)*, Queen's University, Kingston, Canada, August 15-16, 1996, pp.61-76.

[17] J. Moore and G. Simmons, "Cycle Structure of the DES with Weak and Semi-Weak Keys", in *Advances in Cryptology: Proceedings of Crypto '86*, Springer-Verlag, 1987, pp.9-32.

[18] J. Moore and G. Simmons, "Cycle Structure of the DES for Keys Having Palindromic (or Antipalindromic) Sequences of Round Keys", *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, 1987, pp.262-273.

[19] A. Webster, "Plaintext/Ciphertext Bit Dependencies in Cryptographic Systems", *M.Sc. Thesis*, Department of Electrical Engineering, Queen's University, Kingston, Ontario, Canada, 1985.

[20] A. Webster and S. Tavares, "On the Design of S-Boxes", in *Advances in Cryptology: Proceedings of Crypto '85*, Springer-Verlag, 1986, pp.523-534.

Differential cryptanalysis of Feistel ciphers and differentially δ-uniform mappings

Anne Canteaut*
INRIA Projet CODES
Domaine de Voluceau - BP 105
78153 Le Chesnay Cedex - France

Abstract

In this paper we study the round permutations (or S-boxes) which provide to Feistel ciphers the best resistance against differential cryptanalysis. We prove that a Feistel cipher with any round keys and with at least 5 rounds resists any differential attack if its round permutation is differentially δ-uniform for a small δ. This improves an earlier result due to Nyberg and Knudsen which only held for independent and uniformly random round keys. We also give some necessary conditions for a mapping to be almost perfect nonlinear (*i.e.* differentially 2-uniform).

1 Introduction

The underlying motivation of this work is the design of a Feistel cipher which resists all classical attacks. The DES cipher seems to have this property since no cryptanalysis is really more efficient than an exhaustive search for the key. But it would be very important to find a new secure DES-like cipher because the size of the secret-key used in DES makes a brute-force attack feasible. The main problem is therefore to replace the S-boxes used in DES with another function which resists both differential and linear cryptanalysis. In this paper we study the round permutations (which play the same role as the S-boxes) which ensure that the corresponding Feistel cipher is secure against differential cryptanalysis.

In [NK93] Nyberg and Knudsen gave a condition under which a Feistel cipher resists differential cryptanalysis "in average". They actually gave an upper bound on the probability of any r-round differential of a Feistel cipher, for $r \geq 3$, but this bound only holds when the round keys are independent and uniformly random. This result does therefore not rule out the existence of some weak round keys for which a differential attack would be feasible. A lower bound on the complexity of a practical differential attack can then only be deduced if it is additionally assumed that the hypothesis of stochastic equivalence [LMM91] is satisfied, *i.e.* if the differentials have roughly the same probabilities for all round keys. But we here show that this further assumption does usually not

*On leave at Institute for Signal and Information Processing, ETH Zürich, Switzerland

hold for a Feistel cipher. We nevertheless prove that the bound given by Nyberg and Knudsen still holds for any round keys. This stronger result implies that a Feistel cipher resists any differential attack if the round permutation is differentially δ-uniform for a small δ. The resistance of a Feistel cipher against differential cryptanalysis does therefore not require any further assumption on the round keys or on the key scheduling algorithm.

We first briefly recall in Section 2 how differential cryptanalysis works. Section 3 is then devoted to the complexity of a differential attack of a Feistel cipher: we show why Nyberg-Knudsen's result does not suffice to ensure that some Feistel ciphers are practically secure against differential cryptanalysis. We afterwards improve this result since we show that a Feistel cipher with any round keys resists differential cryptanalysis as far as its round permutation is differentially δ-uniform for a small δ. Section 4 gives some general properties of differentially δ-uniform mappings and some necessary conditions for a permutation to be almost perfect nonlinear (APN), *i.e.* differentially 2-uniform. Following a result due to Carlet, Charpin and Zinoviev [CCZ97] we also prove that the smallest value of δ for which a power polynomial is differentially δ-uniform is strongly related to the number of codewords of Hamming weight 3 and 4 in some binary cyclic codes with 2 zeroes.

2 Differential cryptanalysis of iterated ciphers

In an iterated block cipher with r rounds the ciphertext is obtained by iterating r times an invertible function F, called the round function, depending on a secret parameter K called the round key. The r round keys are usually obtained from a unique secret key by a key scheduling algorithm.

A differential attack [BS91] of such an iterated cipher consists in encrypting some plaintexts which only differ from a fixed value α. The difference between two plaintexts X and X' is here defined by a group operation $\otimes$ on the set of plaintexts:

$$\Delta X = X \otimes X'^{-1}$$

where X'^{-1} denotes the inverse of X' with respect to the group operation $\otimes$.

This attack exploits the fact that the round function of an iterated cipher is usually cryptographically weak. This means that the value of the round key K can usually be determined from the knowledge of the difference between the inputs of the function, ΔX, and from both outputs Y and Y'. The basic idea of a differential attack therefore consists in submitting two different plaintexts X and $X' = X \otimes \alpha$ for encryption and in estimating the value of the input difference of the last round $\Delta Y(r-1)$ (see Figure 1). If the round function is cryptographically weak, it is then possible to recover the value of the last-round key K_r.

Differential cryptanalysis will then be successful if there exists an $(r-1)$-*round differential* (α, β) such that

$$P = P[\Delta Y(r-1) = \beta | \Delta Y(0) = \alpha, K_1 = k_1, \ldots, K_{r-1} = k_{r-1}] \qquad (1)$$

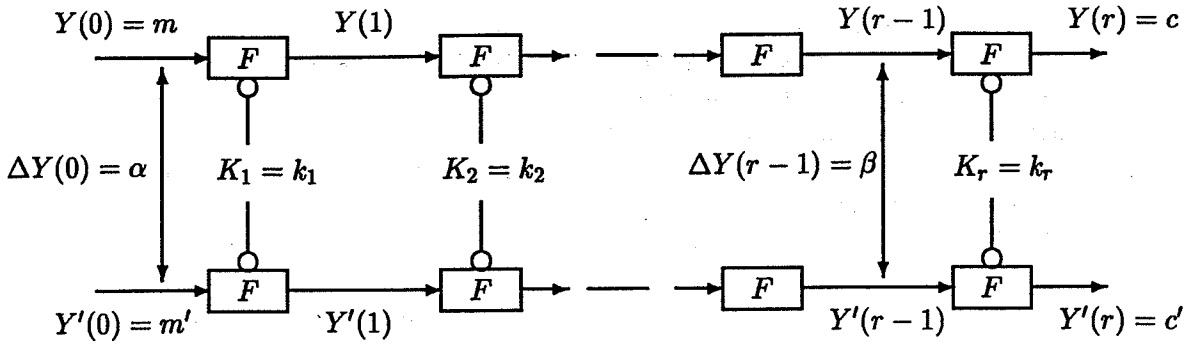


Figure 1: Differential cryptanalysis of an iterated cipher

is high. As soon as such an $(r-1)$-round differential is known, the attack consists in iterating the following procedure:

- Choose a plaintext m uniformly at random and submit m and $m \otimes \alpha$ for encryption.
- Suppose that $\Delta Y(r-1) = \beta$ and determine all corresponding possible values for K_r.

After many steps one value for K_r will occur significantly more often than the other ones. The number of such iterations required for recovering the value of the last-round key is then at least [LMM91]

$$\frac{1}{P - \frac{1}{2^n - 1}}$$

where P is given by Equation (1) and n is the plaintext size.. An iterated cipher then resists differential cryptanalysis if, for a fixed plaintext difference, the probability distribution of the output difference at the last-but-one round is close to the uniform distribution.

The main problem in this attack is to estimate the probability of a differential as expressed in Equation (1) since the first $(r-1)$ round keys are unknown. In most cases we are actually only able to compute the probability of a differential when the round keys are independent and uniformly random, *i.e.* $P[\Delta Y(r-1) = \beta | \Delta Y(0) = \alpha]$. If we want to deduce from this probability whether a differential attack is feasible, we have to assume that the probability of a differential is roughly the same for almost all round keys. This additional condition called *the hypothesis of stochastic equivalence* was pointed out by Lai, Massey and Murphy [LMM91].

Definition 1 *(Hypothesis of stochastic equivalence) For an $(r-1)$-round differential (α, β),*

$$P[\Delta Y(r-1) = \beta | \Delta Y(0) = \alpha, K_1 = k_1, \ldots, K_{r-1} = k_{r-1}] \simeq P[\Delta Y(r-1) = \beta | \Delta Y(0) = \alpha]$$

for almost all round keys $k_1, \ldots k_{r-1}$.

If this hypothesis is not satisfied, a differential may have a low probability in average but its probability may nevertheless be high for some particular round keys. This would mean that some round keys would be weak in the sense that the corresponding cipher would not resist differential cryptanalysis.

3 Resistance of Feistel ciphers against differential cryptanalysis

We are now interested in the complexity of a differential attack of a Feistel cipher when the difference is defined by the bitwise XOR denoted by $+$.

3.1 An upper bound on the average probability of any differential

We here only consider Feistel ciphers with block size $2n$ without expansion. In this case, the round permutation F is designed as follows:

$$\begin{array}{rccc} F\colon & \mathbf{F}_2^n \times \mathbf{F}_2^n & \rightarrow & \mathbf{F}_2^n \times \mathbf{F}_2^n \\ & (L,R) & \mapsto & (R, L + f(R + K_i)) \end{array}$$

where $+$ denotes the exclusive-or operation, $K_i \in \mathbf{F}_2^n$ is the i-th round key and f is a permutation over $\mathbf{F}_2^n$, called the round permutation. Using the particular structure of this round function Nyberg and Knudsen [NK93] gave an upper bound on the probability of any r-round differential for $r \geq 3$ *when the round keys are independent and uniformly random.* They actually proved the following result:

Proposition 1 [NK93] *For a Feistel cipher with block size $2n$, with round permutation f and with independent uniformly random round keys, the probability of any r-round differential (α, β), $\alpha \neq 0$, for $r \geq 3$ satisfies*

$$P[\Delta Y(r) = \beta | \Delta Y(0) = \alpha] \leq \frac{\delta_f^2}{2^{2n}}$$

$$\text{where } \delta_f = \max_{\beta} \max_{\alpha \neq 0} |\{X \in \mathbf{F}_2^n,\ f(X + \alpha) + f(X) = \beta\}|$$

This proposition then implies that any Feistel cipher with at least 5 rounds resists differential cryptanalysis if the round permutation f is such that δ_f is small and if the hypothesis of stochastic equivalence is satisfied. In order to use this theoretical result in practice, Knudsen [Knu94] called a Feistel cipher a *practically secure Feistel cipher* if it resists differential cryptanalysis under the assumption of independent uniformly random round keys. But it unfortunately seems that the hypothesis of stochastic equivalence does not hold in general for a Feistel cipher.

3.2 Hypothesis of stochastic equivalence for Feistel ciphers

As an example we here show that the hypothesis of stochastic equivalence is not satisfied for a small Feistel cipher with block size 8. The round permutation f of this cipher is defined by

$$\begin{array}{rccc} f: & \mathbf{F}_{2^4} & \rightarrow & \mathbf{F}_{2^4} \\ & x & \mapsto & x^7 \end{array}$$

where the vector space $\mathbf{F}_2^4$ is identified with the finite field with 16 elements. For this small Feistel cipher, we give the probabilities of two different 3-round differentials (α, β)

- $\alpha = 00000001$ and $\beta = 011000001$. When the round keys are independent and uniformly random, we obtain

$$P[\Delta Y(3) = \beta | \Delta Y(0) = \alpha] = 9.8\ 10^{-3}$$

  But when the first 3 round keys are fixed, we get

$$\begin{array}{lll} P[\Delta Y(3) = \beta | \Delta Y(0) = \alpha, K_1 = k_1, K_2 = k_2, K_3 = k_3] & = 0 & \text{for 50 \% of the keys} \\ & = 7.8\ 10^{-3} & \text{for 25 \% of the keys} \\ & = 3.12\ 10^{-2} & \text{for 25 \% of the keys} \end{array}$$

- $\alpha = 00010110$ and $\beta = 00010110$. When the round keys are independent and uniformly random, this 3-round differential has probability

$$P[\Delta Y(3) = \beta | \Delta Y(0) = \alpha] = 1.56\ 10^{-2}$$

  But for fixed round keys this probability actually equals

$$\begin{array}{lll} P[\Delta Y(3) = \beta | \Delta Y(0) = \alpha, K_1 = k_1, K_2 = k_2, K_3 = k_3] & = 0 & \text{for 75 \% of the keys} \\ & = 6.25\ 10^{-2} & \text{for 25 \% of the keys} \end{array}$$

It then turns out that for this particular Feistel cipher the hypothesis of stochastic equivalence does not hold. Furthermore the computation of the probability of some 3-round differentials for many different small Feistel ciphers leads to similar results. This then implies that the result given by Nyberg and Knudsen does not allow to deduce if a differential attack of a Feistel cipher is feasible in practice.

3.3 A practical result on the resistance of Feistel ciphers against differential cryptanalysis

The hypothesis of stochastic equivalence is nevertheless satisfied for any Feistel cipher in some particular cases. We here denote by $\delta_f(\alpha, \beta)$ the number of solutions $X \in \mathbf{F}_2^n$ of the equation

$$f(X + \alpha) + f(X) = \beta$$

Proposition 2 *For any Feistel cipher with block size $2n$, the hypothesis of stochastic equivalence exactly holds for any 2-round differential (α, β). Moreover we have for any $\alpha_L, \alpha_R, \beta_L, \beta_R \in \mathbf{F}_2^n$ and for any round keys k_1 and k_2,*

$$P[\Delta Y(2) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2] = P[\Delta Y(2) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R)] = \frac{\delta_f(\alpha_R, \beta_L + \alpha_L)\delta_f(\beta_L, \beta_R + \alpha_R)}{2^{2n}}$$

Proof. We denote by $R(i)$ the right half of the input of the $(i+1)$-th round. Similarly $Z(i) = f(R(i) + K_{i+1})$. When the round keys are fixed, the probability of a 2-round differential can be decomposed as follows:

$$\begin{aligned} P &= P[\Delta Y(2) = (\beta_L, \beta_R) | \Delta Y(0) = \alpha, K_1 = k_1, K_2 = k_2] \\ &= P[\Delta Z(1) = \beta_R + \alpha_R | \Delta R(1) = \beta_L, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2] \\ &\quad \times P[\Delta Z(0) = \beta_L + \alpha_L | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1] \end{aligned}$$

Since $R(0)$ is uniformly random, we obviously have that

$$P[\Delta Z(0) = \beta_L + \alpha_L | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1] = \frac{\delta_f(\alpha_R, \alpha_L + \beta_L)}{2^n}$$

On the other hand, we have

$$\begin{aligned} P[\Delta Z(1) = \beta_R + \alpha_R | \Delta R(1) = \beta_L, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2] = \\ \sum_r (P[\Delta Z(1) = \beta_R + \alpha_R | \Delta R(1) = \beta_L, R(1) + k_2 = r] \\ \times P[R(1) + k_2 = r | \Delta R(1) = \beta_L, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1]) \end{aligned}$$

Since $R(1) = f(R(0) + k_1) + L(0)$ and since $L(0)$ is uniformly distributed, the random variable $R(1)$ is uniformly distributed even if $\Delta R(1)$ and $\Delta Y(0)$ are fixed. We then obtain that

$$P[\Delta Z(1) = \beta_R + \alpha_R | \Delta R(1) = \beta_L, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2] = P[\Delta Z(1) = \beta_R + \alpha_R | \Delta R(1) = \beta_L] = \frac{\delta_f(\beta_L, \alpha_R + \beta_R)}{2^n}$$

□

The hypothesis of stochastic equivalence is also satisfied for some 3-round differentials as asserted in the following proposition.

Proposition 3 *For any Feistel cipher with block size $2n$, the hypothesis of stochastic equivalence exactly holds for any 3-round differential $((\alpha_L, \alpha_R), (\beta_L, \beta_R))$ such that $\alpha_R = 0$ or $\beta_L = \alpha_R$.*

We additionally have that for any round keys k_1, k_2 and k_3,

$$P[\Delta Y(3) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, 0), (K_1, K_2, K_3) = (k_1, k_2, k_3)] = \frac{\delta_f(\alpha_L, \beta_L)\ \delta_f(\beta_L, \beta_R + \alpha_L)}{2^{2n}}$$

$$P[\Delta Y(3) = (\alpha_R, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), (K_1, K_2, K_3) = (k_1, k_2, k_3)] = \frac{\delta_f(\alpha_R, \alpha_L)\ \delta_f(\alpha_R, \beta_R)}{2^{2n}}$$

Proof.

- $\alpha_R = 0$.
  In this case, the first round of the cipher is a trivial round. Thus $\Delta R(1) = \alpha_L$ with probability 1. The random variable $R(1)$ is then uniformly distributed when $L(0)$ and $R(0)$ are uniformly random. We then obtain

$$P[\Delta Y(3) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, 0), (K_1, K_2, K_3) = (k_1, k_2, k_3)] =$$
$$P[\Delta Y(3) = (\beta_L, \beta_R) | \Delta Y(1) = (\alpha_R, \alpha_L), (K_2, K_3) = (k_2, k_3)] = \frac{\delta_f(\alpha_L, \beta_L)\, \delta_f(\beta_L, \beta_R + \alpha_L)}{2^{2n}}$$

  where the last equality is deduced from Proposition 2.

- $\beta_L = \alpha_R$
  In this case $\Delta Z(1) = 0$. Since f is a permutation, this can only occur when $\Delta R(1) = 0$. This implies that the second round of the cipher is here a trivial round. We then have

$$P[\Delta Y(3) = (\alpha_R, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] =$$
$$\frac{\delta_f(\alpha_R, \alpha_L)}{2^n} P[\Delta Z(2) = \beta_R | \Delta R(2) = \alpha_R, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3]$$

  On the other hand the random variable $R(2)$ is uniformly distributed in this case even if the differences $\Delta R(2)$, $\Delta Y(0)$ and the first two round keys are fixed. This implies that

$$P[\Delta Z(2) = \beta_R | \Delta R(2) = \alpha_R, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] =$$
$$P[\Delta Z(2) = \beta_R | \Delta R(2) = \alpha_R]$$

  and we therefore conclude that

$$P[\Delta Y(3) = (\alpha_R, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), (K_1, K_2, K_3) = (k_1, k_2, k_3)] = \frac{\delta_f(\alpha_R, \alpha_L)\, \delta_f(\alpha_R, \beta_R)}{2^{2n}}$$

□

Using that the hypothesis is always satisfied in these both cases, we now prove that the upper bound on the probability of a differential given by Nyberg and Knudsen still holds for any round keys.

Theorem 1 *For a Feistel cipher with block size $2n$, with round permutation f and with any round keys $k_1, \ldots, k_r$, the probability of any r-round differential (α, β), $\alpha \neq 0$, for $r \geq 3$, satisfies*

$$P[\Delta Y(r) = \beta | \Delta Y(0) = \alpha, K_1 = k_1, \ldots, K_r = k_r] \leq \frac{\delta_f^2}{2^{2n}}$$

where $\delta_f = \max_{\beta} \max_{\alpha \neq 0} |\{X \in \mathbf{F}_{2^n},\ f(X + \alpha) + f(X) = \beta\}|$

Proof. We first prove this result for any 3-round differential $((\alpha_L, \alpha_R), (\beta_L, \beta_R))$ with $(\alpha_L, \alpha_R) \neq (0,0)$. The probability of any 3-round differential (α, β) can be decomposed as follows:

$$\begin{aligned} P &= P[\Delta Y(3) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] \\ &= \sum_d P[\Delta Z(2) = \beta_R + d | \Delta R(2) = \beta_L, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] \\ &\quad \times \frac{\delta_f(\alpha_R, d + \alpha_L)\, \delta_f(d, \beta_L + \alpha_R)}{2^{2n}} \end{aligned}$$

- If $\alpha_R \neq \alpha_L$, $\Delta R(1)$ cannot be zero.
  If $\alpha_R \neq 0$, we conclude that

$$\begin{aligned} &P[\Delta Y(3) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] \\ &\leq \frac{\delta_f^2}{2^{2n}} \sum_{d \neq 0} P[\Delta Z(2) = \beta_R + d | \Delta R(2) = \beta_L, \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] \\ &\leq \frac{\delta_f^2}{2^{2n}} \end{aligned}$$

  If $\alpha_R = 0$, the previous proposition gives

$$P[\Delta Y(3) = (\beta_L, \beta_R) | \Delta Y(0) = (\alpha_L, 0), K_1 = k_1, K_2 = k_2, K_3 = k_3] = \frac{\delta_f(\alpha_L, \beta_L)\, \delta_f(\beta_L, \beta_R + \alpha_L)}{2^{2n}} \leq \frac{\delta_f^2}{2^{2n}}$$

  since $\beta_L = 0$ would imply that $\alpha_L = 0$ and hence that $\alpha = 0$.

- If $\alpha_R = \alpha_L$, the previous proposition gives

$$P[\Delta Y(3) = (\alpha_R, \beta_R) | \Delta Y(0) = (\alpha_L, \alpha_R), K_1 = k_1, K_2 = k_2, K_3 = k_3] = \frac{\delta_f(\alpha_R, \alpha_L)\, \delta_f(\alpha_R, \beta_R)}{2^{2n}} \leq \frac{\delta_f^2}{2^{2n}}$$

  since $\alpha_R = 0$ would imply that $\alpha_L = 0$.

We now obtain the same upper bound for any r-round differential for $r > 3$ by induction on r. □

This new theorem implies that a Feistel cipher with any round keys is secure against differential cryptanalysis as far as the round permutation f is such that δ_f is small. This only depends on the following property of the round permutation defined by Nyberg and Knudsen [NK93]:

Definition 2 *A function f over $\mathbf{F}_2^n$ is* differentially δ-uniform *if, for all $\alpha \in \mathbf{F}_2^n$, $\alpha \neq 0$, and for all $\beta \in \mathbf{F}_2^n$,*

$$|\{X \in \mathbf{F}_2^n,\ f(X + \alpha) + f(X) = \beta\}| \leq \delta$$

Using [LMM91, Theorem 1] we obtain a lower bound on the complexity of a differential attack of a Feistel cipher, *i.e.* the number of encryptions it requires.

Corollary 1 *Let us consider a Feistel cipher with block size $2n$, with at least 5 rounds and with a differentially δ-uniform round permutation f. The complexity of a differential attack against this cipher is at least $\frac{2(2^{2n}-1)}{\delta^2-1}$.*

4 Differentially δ-uniform permutations

The number of solutions of $f(X+\alpha)+f(X)=\beta$ is obviously even. This implies that the smallest possible value δ such that a permutation is differentially δ-uniform is $\delta=2$. Differentially 2-uniform permutations are also called *almost perfect nonlinear* (APN) permutations. They correspond to the round permutations which provide the best resistance against differential cryptanalysis.

4.1 APN permutations over $\mathbf{F}_{2^n}$ for even n

From now on we identify the vector-space $\mathbf{F}_2^n$ with the finite field $\mathbf{F}_{2^n}$. Any permutation of $\mathbf{F}_{2^n}$ can be expressed as a unique polynomial of $\mathbf{F}_{2^n}[X]$ of degree at most 2^n-1.

We first give a necessary condition for a polynomial to be APN when n is even.

Proposition 4 *Let n be an even integer. The mapping $f: x \mapsto \sum_{i=0}^{2^n-1} a_i X^i$ is not APN over $\mathbf{F}_{2^n}$ if*

$$\sum_{j=1}^{\frac{2^n-1}{3}} a_{3j} = 0$$

Proof. We first notice that 0 and 1 are two solutions of Equation

$$f(X+1)+f(X) = \sum_{i=0}^{2^n-1} a_i \tag{2}$$

Let now $x=\alpha^u$ where α is a primitive element in $\mathbf{F}_{2^n}$ and $u=\frac{2^n-1}{3}$. Since $x^4=x$, x is in $\mathbf{F}_4$ and $x \notin \{0,1\}$. It then satisfies $x^2+x+1=0$. We then obtain that if $\sum_{j=1}^{\frac{2^n-1}{3}} a_{3j}=0$, $x=\alpha^u$ is another solution of Equation 2. □

This result notably implies that no power polynomial permutation, *i.e.* $f(x)=x^t$ with $\gcd(t,2^n-1)=1$, is APN when n is even.

4.2 Differentially δ-uniform power polynomials and cyclic codes with two zeroes

We now only consider the mappings on $\mathbf{F}_{2^n}$ which can be expressed as a power polynomial X^t. In this case we only have to examine for t a representative of each cyclotomic coset modulo 2^n-1. When f is a power polynomial, the differentially δ-uniform property can be characterized as follows:

Proposition 5 *The power polynomial mapping $f : x \mapsto x^t$ is differentially δ-uniform if and only if for all $c \in \mathbf{F}_{2^n}$, $c \neq 0$, the equation $(X+1)^t + X^t = c$ has at most δ solutions in $\mathbf{F}_{2^n}$.*

In [CCZ97] it is proved that the power polynomial function $x \mapsto x^t$ is APN over $\mathbf{F}_{2^n}$ if and only if the cyclic code $\mathcal{C}_{1,t}$ of length $2^n - 1$ with defining set $\{1, t\}$ has minimum distance 5. The link between differentially δ-uniform power polynomials and cyclic codes is still tighter since the number of solutions of the equations $(X+1)^t + X^t = c$ is related to the number of codewords of weight 3 and 4 in $\mathcal{C}_{1,t}$.

Proposition 6 *Let $\mathcal{C}_{1,t}$ denote the cyclic code of length $2^n - 1$ with defining set $\{1, t\}$ and let δ_c be the number of roots in $\mathbf{F}_{2^n}$ of polynomial $P_c(X) = (X+1)^t + X^t + c$. The number A_3 (resp. A_4) of codewords with Hamming weight 3 (resp. 4) in $\mathcal{C}_{1,t}$ is given by*

$$\begin{aligned} A_3 &= \frac{(2^n-1)}{6}(\delta_1 - 2) \\ A_4 &= \frac{(2^n-1)}{24}\left(\sum_{c \in \mathbf{F}_{2^n}} \delta_c^2 - 2^{n+1} - 4(\delta_1 - 2)\right) \end{aligned}$$

Proof. A binary vector $x = (x_0, \ldots, x_{2^n-2})$ belongs to $\mathcal{C}_{1,t}$ if and only if its syndrome is zero. The word with support $\{i_1, i_2, i_3, i_4\}$ then lies in $\mathcal{C}_{1,t}$ if and only if, for $x_j = \alpha^{ij}$, there exists $(a, b) \in \mathbf{F}_{2^n} \times \mathbf{F}_{2^n}$, $a \neq 0$ such that

$$(x_1 + a)^t + x_1^t = b = (x_3 + a)^t + x_3^t$$

i.e. $x_1 a^{-1}, x_1 a^{-1} + 1, x_3 a^{-1}, x_3 a^{-1} + 1$ are 4 distinct roots of P_c with $c = \frac{b}{a^t}$.

Since 0 is a root of P_c if and only if $c = 1$, we obtain that the codewords of Hamming weight 3 of $\mathcal{C}_{1,t}$ exactly correspond to the 3-tuples $(x, y, x+y)$ with non-zero distinct coordinates such that $x(x+y)^{-1}$ and $x(x+y)^{-1} + 1$ are non-zero roots of P_1. Similarly the codewords of weight 4 in $\mathcal{C}_{1,t}$ exactly correspond to the 4-tuples $(x, y, z, x+y+z)$ with non-zero distinct coordinates such that $x(x+y)^{-1}$, $x(x+y)^{-1} + 1$, $z(x+y)^{-1}$ and $z(x+y)^{-1} + 1$ are 4 distinct roots of P_c. □

Note that if n is odd and if the minimum distance of $\mathcal{C}_{1,t}$ is 3, the smallest possible value for δ such that $x \mapsto x^t$ is differentially δ-uniform is 8 since $\delta_1 \equiv 2 \bmod 6$. Some cyclic codes with 2 zeroes and with minimum distance 3 were examined in [CTZ97].

4.3 Some APN power polynomials

Table 1 lists all known exponents t (up to equivalence) such that $x \mapsto x^t$ is APN. But the only APN power polynomials X^t amongst these 4 families which can be used as a round permutation of a Feistel cipher are those corresponding to $t \in \mathcal{K}_i$ with $i \geq 6$. It was actually proved that the mapping $x \mapsto x^t$ with $t \in \mathcal{I}$ is not secure against linear cryptanalysis [LW90, CV95]. The power polynomials corresponding to $t \in \mathcal{Q}_i$ or $t \in \mathcal{W}$ can neither be used since a differential

exponent	smallest value of δ such that f is differentially δ-uniform	notation for the corresponding cyclotomic coset	ref.
2^i+1	$2^{\gcd(n,i)}$	$\mathcal{Q}_i$	[Nyb93]
2^n-2^i-1	2 if n is odd 4 if n is even	$\mathcal{I}$	[Nyb93, BD93]
$2^{2i}-2^i+1$	2 if n is odd and $\gcd(n,i)=1$	$\mathcal{K}_i$	[Kas71]
$2^{\frac{n-1}{2}}+3$	2 if n is odd	$\mathcal{W}$	[Dob]

Table 1: Minimum value of δ for some power polynomials on $\mathbf{F}_{2^n}$.

attack using higher order differentials is feasible when the Hamming weight of t is small [JK97]. This attack exploits the fact that any ciphertext bit can be expressed as a polynomial in all plaintext bits of degree at most $d = w(t)^{r-3}$ where r denotes the number of rounds

4.4 A lower bound on the degree of APN power polynomials over $\mathbf{F}_{2^n}$

Janwa, McGuire and Wilson [JW93, JMW95] proved that for most values of t, the code $C_{1,t}$ of length 2^n-1 does not have minimum distance 5 for infinitely many values of n. Their proof relies on Weil's theorem which gives a lower bound on the number of rational points on an absolutely irreducible curve over $\mathbf{F}_{2^n}$. We here use a similar argument for proving that for a fixed n the mapping $x \mapsto x^t$ is not APN as far as t exceeds a certain value.

Theorem 2 *Suppose that the curve*

$$g_t(X,Y) = \frac{X^t+Y^t+(X+Y+1)^t}{(X+Y)(X+1)(Y+1)}$$

is absolutely irreducible over $\mathbf{F}_2$. *The mapping* $x \mapsto x^t$ *is not APN over* $\mathbf{F}_{2^n}$, $n \geq 5$, *if*

$$t \leq 2^{\frac{n}{4}} + 4.5$$

Janwa, McGuire and Wilson [JMW95] proved that $g_t(X,Y)$ is absolutely irreducible for any $t \equiv 3 \bmod 4$, $t > 3$ and for some values such that $t \equiv 1 \bmod 4$. They actually conjectured that this curve is absolutely irreducible for all values of t except those lying in the cyclotomic cosets $\mathcal{Q}_i$ and $\mathcal{K}_i$(see Table 1). This statement also holds for any $t < 100$. We therefore give in Table 2 some values of t for which $x \mapsto x^t$ is not APN.

5 Concluding remarks

We here proved that a Feistel cipher without expansion with any round keys resists differential cryptanalysis if its round permutation is differentially δ-uniform

n	7	9	11	13	15	17	19	21	23	25
$t_{\min}$	7	9	11	14	17	23	31	42	58	80

Table 2: Bound $t_{\min}$ such that $x \mapsto x^t$ is not APN over $\mathbf{F}_{2^n}$ for all $t \leq t_{\min}$, $t \notin \mathcal{Q}_i \cup \mathcal{K}_i$

for a small δ. But the only (up to equivalence) known APN permutation which can be used in a Feistel cipher is the power polynomial function over $\mathbf{F}_{2^n}$ defined by $x \mapsto x^{2^{2i}-2^i+1}$ where n is odd and $\gcd(n,i) = 1$. It nevertheless appears that any new result concerning either the number of roots of polynomials over a finite field or the weight distribution of some cyclic codes would have some important consequences for the design of new provably secure Feistel ciphers. It is however important to note that the resistance of Feistel ciphers against a differential attack is still an open problem when the difference is not defined by the bitwise exclusive-or but by another group operation on the set of plaintexts.

References

[BD93] T. Beth and C. Ding. On almost perfect nonlinear permutations. In *Advances in Cryptology - EUROCRYPT'93*, number 765 in Lecture Notes in Computer Science, pages 65–76. Springer-Verlag, 1993.

[BS91] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, 4(1):3–72, 1991.

[CCZ97] C. Carlet, P. Charpin, and V. Zinoviev. Cyclic codes and permutations suitable for DES-like cryptosystems. In *1997 IEEE Information Theory Workshop*, Norway, July 1997. To be presented.

[CTZ97] P. Charpin, A. Tietäväinen, and V. Zinoviev. On binary cyclic codes with $d = 3$. *Problems of Information Transmission*, 1997. To appear.

[CV95] F. Chabaud and S. Vaudenay. Links between differential and linear cryptanalysis. In *Advances in Cryptology - EUROCRYPT'94*, number 950 in Lecture Notes in Computer Science, pages 356–365. Springer-Verlag, 1995.

[Dob] H. Dobbertin. Private Communication.

[JK97] T. Jakobsen and L.R. Knudsen. The interpolation attack on block ciphers. In *Fast Software Encryption 97*, January 1997.

[JMW95] H. Janwa, G. McGuire, and R.M. Wilson. Double-error correcting cyclic codes and absolutely irreducible polynomials over GF(2). *Journal of Algebra*, (178):665–676, 1995.

[JW93] H. Janwa and R.M. Wilson. Hyperplane sections of Fermat varieties in P^3 in char. 2 and some applications to cyclic codes. In *Applied*

Algebra, Algebraic Algorithms and Error-correcting Codes - Proceedings AAECC-10, number 673 in Lecture Notes in Computer Science, pages 180–194. Springer-Verlag, 1993.

[Kas71] T. Kasami. The weight enumerators for several classes of subcodes of the second order binary Reed-Muller codes. *Information and Control*, (18):369–394, 1971.

[Knu94] L.R. Knudsen. Practically secure Feistel ciphers. In *Fast Software Encryption 93*, number 809 in Lecture Notes in Computer Science, pages 211–221. Springer-Verlag, 1994.

[LMM91] X. Lai, J.L. Massey, and S. Murphy. Markov ciphers and differential cryptanalysis. In *Advances in Cryptology - EUROCRYPT'91*, number 547 in Lecture Notes in Computer Science, pages 17–38. Springer-Verlag, 1991.

[LW90] G. Lachaud and J. Wolfmann. The weights of the orthogonal of the extended quadratic binary Goppa codes. *IEEE Transactions on Information Theory*, 36(3):686–692, 1990.

[NK93] K. Nyberg and L.R. Knudsen. Provable security against differential cryptanalysis. In *Advances in Cryptology - CRYPTO'92*, number 740 in Lecture Notes in Computer Science, pages 566–574. Springer-Verlag, 1993.

[Nyb93] K. Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT'93*, number 765 in Lecture Notes in Computer Science, pages 55–64. Springer-Verlag, 1993.

On Provable Security against Differential and Linear Cryptanalysis in Generalized Feistel Ciphers with Multiple Random Functions

Yasuyoshi KANEKO[†1] Fumihiko SANO[‡2] Kouichi SAKURAI[‡3]

† Telecommunications Advancement Organization of Japan,
Shin-urashima-cho, Kanagawa-ku, Yokohama 221, JAPAN
kaneko@yokohama.tao.or.jp

‡ Dept. of Computer Science and Communication Engineering,
Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812-81, JAPAN
{sano,sakurai}@csce.kyushu-u.ac.jp

Abstract

We investigate the cryptographic role of random functions used in Generalized Feistel Ciphers in achieving provable security against differential and linear cryptanalysis.

The provable security against differential and linear cryptanalysis of block ciphers can be estimated from the maximum probabilities of differential and linear hull. In case of DES-like block ciphers, these probabilities are known to be smaller than twice the square of the maximum differential and linear hull probabilities of one-round. Even though differential characteristic and linear approximation probabilities decrease as the serial round iteration number increases, known upper-bounds of probabilities of differential and linear hull are constant and not less than the square or twice the square of the maximum of the probabilities of one-round. It seems an unproven conjecture that increasing the serial iteration number would fail to achieve stronger provable security against differential and linear cryptanalysis.

This paper introduces the *Generalized Feistel Ciphers*, multiple random functions are used in a usual Feistel network, whereas most DES-like block ciphers use only one random function. We prove that the proposed Generalized Feistel Ciphers achieve the estimation in which the upper bound of the differential probability is strictly less than or equal to the square of the maximum differential probability of one-round, even if a non-injective function is there in the possible position. We also show that a kind of duality holds between the differential probabilities and the linear hull probabilities among these Generalized Feistel Ciphers, which implies that the similar as our obtained results on provable security against differential cryptanalysis holds for provable security against linear cryptanalysis according to these relations of the duality.

Keywords

DES-like block ciphers, Generalized Feistel Ciphers, Differential cryptanalysis, Linear cryptanalysis, Provable security, Design of block ciphers

1 Introduction

This paper investigates what types of block ciphers have provable security against differential cryptanalysis [BS91] and linear cryptanalysis [Mat93] and which types have desirable provable security.

The Data Encryption Standard and Recent Cryptanalysis: The Data Encryption Standard (DES) [NBS77], published in the 1970's, is now the most widely used cipher throughout the world. Due to rapid advances in cryptanalysis as well as computing technology over the past 20 years, particularly the recent discovery of differential cryptanalysis by Biham and Shamir [BS91] and linear cryptanalysis by Matsui [Mat93], the cryptographic strength of DES is being questioned by an increasing number of researchers as well as practitioners. Structurally DES can be viewed as being obtained by the iteration of a basic transform which was first proposed by Feistel [F73, FNS75] and will be also called a DES-like transform in this paper.

Toward the design of provably secure block ciphers: The security of a block cipher against differential cryptanalysis is characterized by the differential characteristic. If we design a block cipher secure against differential attack, the block cipher should have a very small probability of differential characteristic. Furthermore, Lai, Massey, and Murphy [LMM91] observed that we should consider not only the maximum

[1]This study is a part of an activity within the Information & Communication Security Project of Ministry in Japan. He is temporarily transferred from Hitachi, Ltd.

[2]He is now working for TOSHIBA CORPORATION.

[3]He is a sub-leader of the Project.

differential characteristic but also the maximum differential probability for more accurately evaluating the security of block ciphers. We should remark that the maximum differential probability is hard to compute for a given block cipher, though the maximum differential characteristic can be obtained for DES and FEAL by using elegant algorithms [Mat94, OMA95].

Recent work by some researchers, based on the estimation of the maximum differential probability, shows that by using a function secure against differential and linear cryptanalysis, it is possible to construct DES-like block ciphers that are provably secure against differential and linear cryptanalysis [NK95, Nyb94]. Nyberg and Knudsen [NK95] proved that the maximum differential probability (DP_{max}) of Feistel ciphers with 4 rounds can be estimated from the maximum differential probability of one-round(dp_{max}) as $DP_{max} \leq 2dp^2_{max}$. Furthermore, Aoki and Ohta [AO97] showed that, if the used random function is a permutation, then the relation $DP_{max} \leq dp^2_{max}$ holds for Feistel ciphers with 3 rounds. Matsui [Mat96] proposed new types of Feistel ciphers and practical methods for implementing Nyberg et al.'s idea. Recently Nyberg [Nyb96] further considered a design of a block cipher based on a generalized Feistel network originated from Generalized DES [S-B83], which uses smaller s-boxes.

Limitations of the previous approach: A limitation of Nyberg et al.'s approach is that even if we increase the serial-iteration number of the basic DES-like ciphers, we fail to show that the corresponding block ciphers hold a smaller upper bound to maximum differential probability than ciphers with 3 or 4 rounds. The other problem is that results by Aoki and Ohta [AO97] indicate that in case of DES-like block ciphers permutations can achieve more provably secure against differential and linear cryptanalysis than non-injective functions, though designing cryptographic permutations is harder than designing cryptographic non-permutations.

Our results: The contribution of this paper is to show that ciphers more provably secure against differential and linear cryptanalysis can be achieved by using a general Feistel structure with multiple functions per round. Our proposed Generalized Feistel Ciphers(first defined in [Sch93]) are based on the serial iteration of the following one round structure:

$$(Y_H, Y_L) = (S(X_L, k_2), T(X_H, k_1) \oplus f(S(X_L, k_2), k_3)).$$

Note that functions S and T must be permutations because of the requirement of decipherbility, whereas function f is not necessarily a permutation.

We prove that, for some types of Generalized Feistel Ciphers, the relation $DP_{max} \leq dp^2_{max}$ holds for 2 round iterations, even if they adopt a non-injective function. Moreover, we prove that DP_{max} is exactly less than dp^2_{max} by using joined estimation parameters of non-zero minimum differential and linear hull probabilities of function f. We should remark that no block ciphers is known whose DP_{max} is smaller than dp^2_{max}. We also show that a kind of duality between differential probabilities and linear hull probabilities holds among our classified Generalized Feistel Ciphers. This kind of duality was indicated in [Mat94] for DES-like block ciphers with a single function and this duality implies that the similar as our obtained result on provable security against differential cryptanalysis holds for provable security against linear cryptanalysis.

2 Definition

2.1 Generalized Feistel Ciphers

In this paper a r round block cipher is constructed with round function $Y = F(X, K)$ and $X, Y \in GF(2)^N, K \in GF(2)^M$, where X is an input, K is a round key and Y is an output. From a plain text P a cipher text C is obtained as $C = F(\ldots F(F(P, K_1), K_2) \ldots, K_r)$. Where i-th round key $K_i (1 \leq i \leq r)$ is an element of $GF(2)^M$, and each K_i is supposed to be uniformly random. The function F is bijective in terms of input X. When this round function F is as described in Fig. 1 we call the block ciphers consisting of F Generalized Feistel Ciphers [Sch93]. i.e. round function F is constructed with some random functions $y = f(x, k), y = S(x, k)$ or $y = T(x, k)$, where $x, y \in GF(2)^n, k \in GF(2)^m$ and a bitwize XOR operator $\oplus$. In particular, functions S and T are both injective in terms of input x, but function f is not necessarily injective.

Generalized Feistel Ciphers are classified as follows.

1. Basic type (called B-type); round function F is constructed with only one random function f (Fig.2).

2. Left-sided type (called L-type); round function F is constructed with only one injective random function T (Fig.3).

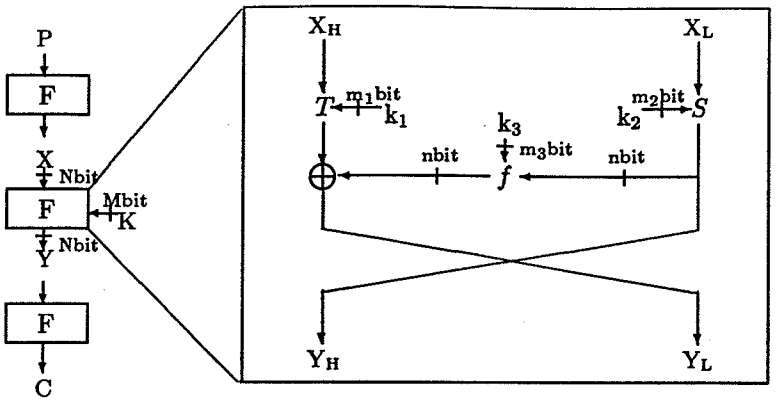


Figure 1: Structure of Generalized Feistel Ciphers

3. Right-sided type (called R-type); round function F is constructed with only one injective random function S (Fig.4).

4. Left-sided complex type (called LB-type); round function F is constructed with random functions f and T (Fig.5).

5. Right-sided complex type (called RB-type); round function F is constructed with random functions f and S (Fig.6).

6. Both-sided type (called LR-type); round function F is constructed with random functions S and T (Fig.7).

7. Both-sided complex type (called LRB-type); round function F is constructed with random functions f, S and T (Fig.8).

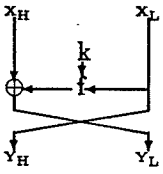


Fig.2 B-type

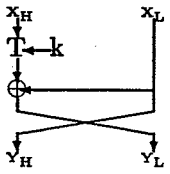


Fig.3 L-type

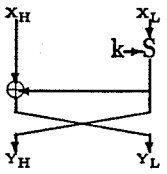


Fig.4 R-type

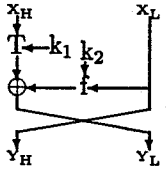


Fig.5 LB-type

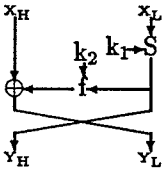


Fig.6 RB-type

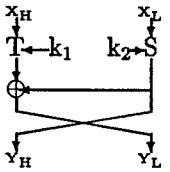


Fig.7 LR-type

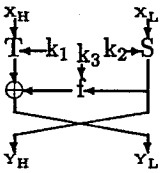


Fig.8 LRB-type

Concerning about key schedules, by modifing the well known key schedule structure of DES, round keys of each types of Generalized Feistel Ciphers are considered to be generated(see APPENDIX A). In this case, the secret key length of each type of Generalized Feistel Ciphers is the same and individual round keys k_1, k_2 and k_3 are generated from different compression permutations.

2.2 Probabilities of r-round Differential and Linear Hull

To analyze r-round iterated block ciphers from the viewpoint of provable security against differential and linear cryptanalysis, we define the differential probability dp of random functions and linear hull probability lp of random functions as follows.

Definition 1 [LMM91], [Mat96], [Nyb94] *Let $y = f(x,k)$ with $x \in GF(2)^n, k \in GF(2)^m$ be a general random function , for given differential values Δx and $\Delta y \in GF(2)^n$ the differential probability of f is*

defined as following.

$$dp_f(\Delta x \to \Delta y) := \underset{k}{aver}\, \underset{x}{Prob}\{f(x \oplus \Delta x, k) \oplus f(x,k) = \Delta y\}.$$

For given masking values Γx *and* $\Gamma y \in GF(2)^n$ *the linear hull probability of* f *is given by*

$$lp_f(\Gamma y \to \Gamma x) := \underset{k}{aver}\, |2\, \underset{x}{Prob}\{x \bullet \Gamma x \oplus f(x,k) \bullet \Gamma y = 0\} - 1|^2.$$

where $\bullet$ *denotes the dot product between a pair of elements in* $GF(2)^n$.

$dp_f(\Delta x \to \Delta y)$ and $lp_f(\Gamma y \to \Gamma x)$ satisfy the followng conditions.

Property 1 (Property of transition probability) [AO97], [Mat96]

$$\sum_{\Delta y} dp_f(\Delta x \to \Delta y) = 1, \qquad \sum_{\Gamma x} lp_f(\Gamma y \to \Gamma x) = 1.$$

When f *is bijective,*

$$\sum_{\Delta x} dp_f(\Delta x \to \Delta y) = 1, \qquad \sum_{\Gamma y} lp_f(\Gamma y \to \Gamma x) = 1,$$

are satisfied.

Property 2 [AO97], [Mat96] f *is injective, if and only if* $dp_f(\Delta x \to 0) = 0$ *for all* $\Delta x \neq 0$ *or* $lp_f(\Gamma y \to 0) = 0$ *for all* $\Gamma y \neq 0$.

Definition 2 [OG94] *Let* χ *be a type of Generalized Feistel Cipher and* r *be the number of it's rounds. For the* i*-th round input* $X(i) = (X_H(i), X_L(i)), 1 \le i \le r$ *and the* i*-th round output* $Y(i) = (Y_H(i), Y_L(i))$, $1 \le i \le r$ *with* $X(1) = (P_H, P_L)$ *and* $Y(r) = (C_H, C_L)$, *the probability of* r*-round differential on* χ *is defined as a Markov chain as follows.*

$$DP_B(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \prod_{i=1}^{r} dp_f\big(\Delta X_L(i) \to \Delta X_H(i) \oplus \Delta Y_L(i)\big).$$

$$DP_L(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \prod_{i=1}^{r} dp_T\big(\Delta X_H(i) \to \Delta X_L(i) \oplus \Delta Y_L(i)\big).$$

$$DP_R(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \prod_{i=1}^{r} dp_S\big(\Delta X_L(i) \to \Delta Y_H(i)\big).$$

$$DP_{LB}(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \sum_{\Delta u} \prod_{i=1}^{r} dp_T\big(\Delta X_H(i) \to \Delta u(i)\big) dp_f\big(\Delta X_L(i) \to \Delta u(i) \oplus \Delta Y_L(i)\big).$$

$$DP_{RB}(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \prod_{i=1}^{r} dp_S\big(\Delta X_L(i) \to \Delta Y_H(i)\big) lp_f\big(\Delta Y_H(i) \to \Delta X_H(i) \oplus \Delta Y_L(i)\big).$$

$$DP_{LR}(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \prod_{i=1}^{r} dp_T\big(\Delta X_H(i) \to \Delta Y_H(i) \oplus \Delta Y_L(i)\big) dp_S\big(\Delta X_L(i) \to \Delta Y_H(i)\big).$$

$$DP_{LRB}(r, \Delta P \to \Delta C) := \sum_{\Delta X, \Delta Y} \sum_{\Delta w} \prod_{i=1}^{r} dp_T\big(\Delta X_H(i) \to \Delta w(i)\big) dp_S\big(\Delta X_L(i) \to \Delta Y_H(i)\big) dp_f\big(\Delta Y_H(i) \to \Delta w(i) \oplus \Delta Y_L(i)\big).$$

Here $\sum_{\Delta X, \Delta Y}$ *denotes the sum total over these parameters as* $\Delta X_H(2), \ldots, \Delta X_H(r)$, $\Delta X_L(2), \ldots, \Delta X_L(r)$, $\Delta Y_H(1), \ldots, \Delta Y_H(r-1)$, $\Delta Y_L(1), \ldots, \Delta Y_L(r-1)$ *and* $\Delta X(i) = \Delta Y(i-1)$ *for all* i. *Each* $\sum_{\Delta u}$ *or* $\sum_{\Delta w}$ *denotes the sum total over these parameters as* $\Delta u(1), \ldots, \Delta u(r)$ *or* $\Delta w(1), \ldots, \Delta w(r)$.

For given masking values ΓP *and* ΓC, *the probability of* r*-round linear hull on* χ *is defined as a Markov chain as follows.* $\sum_{\Gamma X, \Gamma Y}$ *denotes the sum total over these parameters as* $\Gamma X_H(2), \ldots, \Gamma X_H(r)$,

$\Gamma X_L(2),\ldots,\Gamma X_L(r)$, $\Gamma Y_H(1),\ldots,\Gamma Y_H(r-1)$, $\Gamma Y_L(1),\ldots,\Gamma Y_L(r-1)$ *and* $\Gamma X(i)=\Gamma Y(i-1)$ *for all* i. *Each* $\sum_{\Gamma v}$ *or* $\sum_{\Gamma z}$ *denotes the sum total over these parameters as* $\Gamma v(1),\ldots,\Gamma v(r)$ *or* $\Gamma z(1),\ldots,\Gamma z(r)$.

$$LP_B(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\prod_{i=1}^{r} lp_f\big(\Gamma Y_L(i)\rightarrow\Gamma X_L(i)\oplus\Gamma Y_H(i)\big).$$

$$LP_L(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\prod_{i=1}^{r} lp_T\big(\Gamma Y_L(i)\rightarrow\Gamma X_H(i)\big).$$

$$LP_R(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\prod_{i=1}^{r} lp_S\big(\Gamma X_H(i)\oplus\Gamma Y_H(i)\rightarrow\Gamma X_L(i)\big).$$

$$LP_{LB}(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\prod_{i=1}^{r} lp_T\big(\Gamma Y_L(i)\rightarrow\Gamma X_H(i)\big)lp_f\big(\Gamma Y_L(i)\rightarrow\Gamma Y_H(i)\oplus\Gamma X_L(i)\big).$$

$$LP_{RB}(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\sum_{\Gamma v}\prod_{i=1}^{r} lp_S\big(\Gamma v(i)\rightarrow\Gamma X_L(i)\big)lp_f\big(\Gamma X_H(i)\rightarrow\Gamma v(i)\oplus\Gamma Y_H(i)\big).$$

$$LP_{LR}(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\prod_{i=1}^{r} lp_T\big(\Gamma Y_L(i)\rightarrow\Gamma X_H(i)\big)lp_S\big(\Gamma Y_H(i)\oplus\Gamma Y_L(i)\rightarrow\Gamma X_L(i)\big).$$

$$LP_{LRB}(r,\Gamma C\rightarrow\Gamma P) := \sum_{\Gamma X,\Gamma Y}\sum_{\Gamma z}\prod_{i=1}^{r} lp_T\big(\Gamma Y_L(i)\rightarrow\Gamma X_H(i)\big)lp_S\big(\Gamma z(i)\rightarrow\Gamma X_L(i)\big)$$
$$lp_f\big(\Gamma Y_L(i)\rightarrow\Gamma z(i)\oplus\Gamma Y_L(i)\big).$$

For above random functions f, S and T, let p_f, p_S and p_T be maximum values of each $dp_f(\Delta x\rightarrow\Delta y)$, $dp_S(\Delta x\rightarrow\Delta y)$ and $dp_T(\Delta x\rightarrow\Delta y)$ for all such $\Delta x\neq 0$ and Δy. Further, let q_f, q_S and q_T be maximum values of each $lp_f(\Gamma y\rightarrow\Gamma x)$, $lp_S(\Gamma y\rightarrow\Gamma x)$ and $lp_T(\Gamma y\rightarrow\Gamma x)$ for all such $\Gamma y\neq 0$ and Γx. In the case of provable security against differential cryptanalysis, each upper bound to maximum values $DP_\chi(r) := max_{\Delta P\neq 0,\Delta C}DP_\chi(r,\Delta P\rightarrow\Delta C)$ is expressed by polynomials of parameters p_f, p_S or p_T [NK95]. In the case of provable security against linear cryptanalysis, each upper bound to maximum values of $LP_\chi(r) := max_{\Gamma C\neq 0,\Gamma P}LP_\chi(r,\Gamma C\rightarrow\Gamma P)$ is expressed by polynomials of parameters q_f, q_S or q_T [Nyb94].

About these probabilities of differential and linear hull, the following results are easily confirmed from the above definitions.

Property 3 *For any type of* χ

$$DP_\chi(r+1)\leq DP_\chi(r) \quad \textit{for any } r\geq 1,$$
$$LP_\chi(r+1)\leq LP_\chi(r) \quad \textit{for any } r\geq 1.$$

These properties give the expectation that upper bounds to $DP_\chi(r)$ and $LP_\chi(r)$ also decrease as the number of round r increases for any type of Generalized Feistel Ciphers χ. We prove that this expectation is true excluding LR-type in Section 4.

Property 4 *For all types of* χ, $DP_\chi(r,\Delta P\rightarrow\Delta X)$ *and* $LP_\chi(r,\Gamma C\rightarrow\Gamma X)$ *define the transition probabilities as follows.*

$$\sum_{\Delta X} DP_\chi(r,\Delta P\rightarrow\Delta X)=1 \quad \textit{for any } r\geq 1,$$
$$\sum_{\Gamma X} LP_\chi(r,\Gamma C\rightarrow\Gamma X)=1 \quad \textit{for any } r\geq 1.$$

These properties are effectively used in our estimation of upper bounds.

2.3 Duality between Probabilities of Differential and Linear Hull

Between r-round probabilities of differential and linear hull on Generalized Feistel Ciphers, there are relations of duality. This duality is different from the 'links' given by S.Vaudenay et al. [CV94]. In the case of B-type block ciphers a relation of duality has been known as the duality structure between differential cryptanalysis and linear cryptanalysis on DES-like ciphers [AO97], [Mat94].

Theorem 1

$$DP_B(r, \Delta P \to \Delta C) \Leftrightarrow LP_B(r, \Gamma C \to \Gamma P) \quad (1)$$
$$DP_L(r, \Delta P \to \Delta C) \Leftrightarrow LP_R(r, \Gamma C \to \Gamma P) \quad (2)$$
$$DP_R(r, \Delta P \to \Delta C) \Leftrightarrow LP_L(r, \Gamma C \to \Gamma P) \quad (3)$$
$$DP_{LB}(r, \Delta P \to \Delta C) \Leftrightarrow LP_{RB}(r, \Gamma C \to \Gamma P) \quad (4)$$
$$DP_{RB}(r, \Delta P \to \Delta C) \Leftrightarrow LP_{LB}(r, \Gamma C \to \Gamma P) \quad (5)$$
$$DP_{LR}(r, \Delta P \to \Delta C) \Leftrightarrow LP_{LR}(r, \Gamma C \to \Gamma P) \quad (6)$$
$$DP_{LRB}(r, \Delta P \to \Delta C) \Leftrightarrow LP_{LRB}(r, \Gamma C \to \Gamma P) \quad (7)$$

(In this place $\Leftrightarrow$ shows it is possible to change each other)

According to these expressions of $DP_\chi(r, \Delta P \to \Delta C)$ and $LP_\chi(r, \Gamma C \to \Gamma P)$ in Definition 2, above each reciprocal changing $\Leftrightarrow$ is achieved by bijective correspondence between differential values $\Delta X_*(i), \Delta Y_*(i)$ and masking values $\Gamma X_*(i), \Gamma Y_*(i)$.
For example of (2) correspondences $\Delta X_H(i) \leftrightarrow \Gamma X_L(i), \Delta X_L(i) \leftrightarrow \Gamma X_H(i)$ and $\Delta Y_L(i) \leftrightarrow \Gamma Y_H(i)$ bijectively change dp_T to lp_S. In point of changing between dp_T and lp_S(or dp_S and lp_T), the relation between input and output is also changed by each other(this changing is assured by the injectivity of functions S and T). Same correspondences are applied to each cases of (1), (3), (4), (5), (6) and (7).

From these relations of duality, estimation results to r-round differential probabilities are adjusted to correspondence r-round linear hull probabilities with respectively changing p_f, p_S and p_T to q_f, q_T and q_S.

3 Previous Results and Some Conjectures(Provable Security on B-type, L-type and R-type Block Ciphers)

This subsection examines previous estimations of provable security and some conjectures about the estimations are considered.

3.1 Conjectures about Injectivity

3.1.1 B-type

B-type structure is used by DES-like ciphers and was first estimated as following.

Theorem 2 [Nyb94] [NK95] *For any* $r \geq 4$

$$DP_B(r) \leq 2p_f^2 , \qquad LP_B(r) \leq 2q_f^2.$$

These results were obtained from analyzing 4-round DES-like ciphers and using the fact of Property 3. So as mentioned in introduction this estimation result does not depend on the number of round r. But as many experiment results [BS91, Mat93, OMA95] and thoretical results [LMM91, OG94] show that Feistel ciphers are possible to be cryptanalytically strengthened by increasing these iterating number, so in this cases of estimation these upper bounds to $DP_B(r)$ and $LP_B(r)$ are also expected to change depending on the number of round r.

Conjecture 1 *Upper bounds to probabilities of r-round differential and r-round linear hull decrease as the number of rounds increases.*

The following subsection introduces new estimation parameters to confirm that this conjecture is true in cases of Generalized Feistel Ciphers having random function f. On the other hand, it was shown that if random function f is injective, the upper bounds become more small as follows.

Theorem 3 [AO97] *If f is a permutation then*

$$DP_B(r) \leq p_f^2, \qquad LP_B(r) \leq q_f^2 \qquad \text{for any } r \geq 3.$$

This estimation is obtained from essentially using Property 2. The converse statement of this theorem is not yet confirmed.

Conjecture 2 *If f is noninjective, upper bounds to probabilities of r-round differential and r-round linear hull are no less than p_f^2 and q_f^2.*

However, the next subsection proves that without this injectivity assumption for random function f, all Feistel ciphers of LB-type have the same upper bound of p_f^2.

3.2 Conjectures for Lower Limit of Upper Bound

3.2.1 L-type

The L-type structure is used by the new block cipher MISTY2 and has been estimated as follows.

Theorem 4 [Mat96] *For any* $r \geq 3$

$$DP_L(r) \leq {p_T}^2 , \qquad LP_L(r) \leq {q_T}^2.$$

3.2.2 R-type

R-type has not been used yet but from Theorem 4 above and the duality property of Theorems 1 (2) and (3), an estimation is easily achieved as follows.

Theorem 5 *For any* $r \geq 3$

$$DP_R(r) \leq p_S^2 , \qquad LP_R(r) \leq q_S^2.$$

In any cases of these simple types of B, L and R, provable security against differential and linear crytpt-analysis seems to depend on whether the random functions used are injective or not.

Conjecture 3 *Upper bounds to the probabilities of differential and linear hull are not less than the square of* p_* *and* q_*.

The following subsection introduces new estimation parameters to confirm that this conjecture fails in cases of LB and RB-types.

4 New Results (Provable Security on LB-type, RB-type, LR-type and LRB-type Block Ciphers)

4.1 Estimation of Provable Security to Differential Probability of LB-type

This subsection examines The provable security of LB-type.

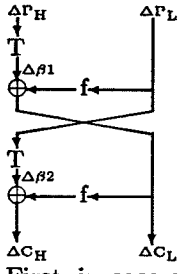
Theorem 6 *The upper bound to the* r*-round differential probability of LB-type is estimated as*

$$DP_{LB}(r) \leq max\{p_T p_f, p_f^2\} \quad \textit{for any } r \geq 2.$$

Here injectivity of f is not assumed, but when $p_T \leq p_f$ it happens that $DP_{LB}(r) \leq p_f^2$(for any $r \geq 2$). This is negative answer to Conjecture 2.

Proof

Property 3 is enough to prove the case of $r = 2$.



$$\begin{aligned} & DP_{LB}(2, \Delta P \rightarrow \Delta C) \\ &= \sum_{\Delta\beta 1, \Delta\beta 2} dp_T(\Delta P_H \rightarrow \Delta\beta 1) dp_f(\Delta P_L \rightarrow \Delta\beta 1 \oplus \Delta C_L) \\ & \qquad dp_T(\Delta P_L \rightarrow \Delta\beta 2) dp_f(\Delta C_L \rightarrow \Delta C_H \oplus \Delta\beta 2). \end{aligned}$$

First, in case of $\Delta C_L = 0$, as $\Delta C \neq 0$ from the assumption $\Delta P \neq 0$, last terms of above equation give $\Delta\beta 2 = \Delta C_H \neq 0$.

Moreover, when $\Delta P_L = 0$, as $dp_T(\Delta P_L \rightarrow \Delta\beta 2) = 0$ (by Property 2), the estimation becomes

$$DP_{LB}(2, \Delta P \rightarrow \Delta C) = 0.$$

Elesewhere when $\Delta P_L \neq 0$, as $dp_T(\Delta P_L \rightarrow \Delta\beta 2) \leq p_T$, the estimation becomes

$$\begin{aligned} DP_{LB}(2, \Delta P \rightarrow \Delta C) &\leq p_T \sum_{\Delta\beta 1} dp_T(\Delta P_H \rightarrow \Delta\beta 1) dp_f(\Delta P_L \rightarrow \Delta\beta 1 \oplus \Delta C_L) \\ &\leq p_T p_f \sum_{\Delta\beta 1} dp_T(\Delta P_H \rightarrow \Delta\beta 1) \\ &= p_T p_f. \qquad \text{(by Property 1)} \end{aligned}$$

Second, in the case of $\Delta C_L \neq 0$,

$$\begin{aligned} DP_{LB}(2, \Delta P \to \Delta C) &\leq p_f \sum_{\Delta\beta 1, \Delta\beta 2} dp_T(\Delta P_H \to \Delta\beta 1) dp_f(\Delta P_L \to \Delta\beta 1 \oplus \Delta C_L) dp_T(\Delta P_L \to \Delta\beta 2) \\ &= p_f \sum_{\Delta\beta 1} dp_T(\Delta P_H \to \Delta\beta 1) dp_f(\Delta P_L \to \Delta\beta 1 \oplus \Delta C_L). \end{aligned}$$

Thus when $\Delta P_L = 0$, with paying attention to $\Delta\beta 1 = \Delta C_L \neq 0$ and $\Delta P_H \neq 0$,

$$DP_{LB}(2, \Delta P \to \Delta C) \leq p_f dp_T(\Delta P_H \to \Delta C_L) \leq p_T p_f.$$

Last when $\Delta P_L \neq 0$,

$$DP_{LB}(2, \Delta P \to \Delta C) \leq {p_f}^2 \sum_{\Delta\beta 1} dp_T(\Delta P_H \to \Delta\beta 1) = {p_f}^2.$$

From the above results we obtain the estimation;

$$DP_{LB}(2) = max\{p_T p_f, {p_f}^2\}.$$

(Q.E.D.)

This LB-type of Generalized Feistel Ciphers has a round that uses 2 times as much key as the Feistel ciphers such as B-type, L-type and R-type. So 2-round of this LB-type could be compared to a 4-round one of the usual type, but as it has shown in these Theorem 2, 3, 4 and 5, estimations of these 4-round usual Feistel ciphers are $2p_f^2$(in case of B-type, here injectivity of f is not assumed) and p_T^2(in case of L). If LB-type is considered to be obtained by strengthening the simple B-type by adding a more secure function T, then as it is usual to suppose $p_T \leq p_f$ so $max\{p_T p_f, p_f^2\} \leq 2p_f^2$. If LB-type is considered to be obtained by strengthening the simple L-type by adding a more secure function f, then as $p_f \leq p_T$ so max $\{p_T p_f, p_f^2\} \leq p_T^2$. These strengthening processes are summarized as $max\{p_T p_f, p_f^2\} \leq max\{2p_f^2, p_T^2\}$, and LB-type seems more secure Feistel ciphers against differential and linear cryptanalysis than L-type and B-type.

Moreover, when adding non-zero minimum differential probability of f defined as

$$\varepsilon_f := min_{\Delta \neq 0, \Delta y} dp_f(\Delta x \to \Delta y) > 0$$

to the estimation parameters, the upper bound is estimated more small as follows.

Theorem 7 *If $p_f \geq p_T$ is satisfied then following estimations are possible.*

$$\begin{aligned} DP_{LB}(2k) &\leq max\{ \; p_T p_f - (1-p_T) p_T \textstyle\sum_{i=1}^{k-1} {\varepsilon_f}^i {p_f}^i, \\ &\qquad {p_f}^2 - (1-p_f)(p_f - p_T) \textstyle\sum_{i=1}^{k-1} {\varepsilon_f}^i {p_f}^i\} \qquad \textit{for any } \; k \geq 1, \\ DP_{LB}(2k-1) &\leq max\{ \; p_T p_f - (1-p_f) p_T \textstyle\sum_{i=1}^{k-2} {\varepsilon_f}^i {p_f}^i, \\ &\qquad {p_f}^2 - (1-p_f)(p_f - p_T) \textstyle\sum_{i=1}^{k-2} {\varepsilon_f}^i {p_f}^i\} \qquad \textit{for any } k \geq 2. \end{aligned}$$

This estimation means that provable security depends on round number and as round number increases the probabilities of upper bounds decrease. This gives an affirmative answer to Conjecture 1. Moreover, as it happens that $DP_{LB}(r)$ is actually smaller than p_f^2, we can get a negative answer to Conjecture 3. Especially the limit of this upper bound gives a lowest upper bound as

$$max\{p_T p_f - \frac{1-p_T}{1-\varepsilon_f p_f} \varepsilon_f p_f p_T, p_f^2 - \frac{(1-p_f)(p_f - p_T)}{1-\varepsilon_f p_f} \varepsilon_f p_f\}.$$

$\frac{1-p_T}{1-\varepsilon_f p_f} \varepsilon_f p_f p_T$ or $\frac{(1-p_f)(p_f - p_T)}{1-\varepsilon_f p_f} \varepsilon_f p_f$ is considered to be a gap between upper bounds of Theorem 6 and Theorem 7. When defining t as $t := p_f - \varepsilon_f$, this gap becomes a decrease function of a parameter t and this fact shows that narrow distribution of differential probabilities of one-round may be effective to strengthen the Feistel ciphers. The proof of this theorem is given by following step.

Lemma 1 *Given a differential value of plaintext $\Delta P = (\Delta P_H, \Delta P_L)$ and a differential value of ciphertext $\Delta C = (\Delta C_H, \Delta C_L)$,*
when $\Delta C_L = 0$

$$DP_{LB}(r, \Delta P \to \Delta C)|_{(\Delta C_L = 0)} \leq p_T p_f - p_T p_f A(r-2), \tag{1}$$

and when $\Delta C_L \neq 0$

$$DP_{LB}(r, \Delta P \to \Delta C)|_{(\Delta C_L \neq 0)} \leq p_f^2 - p_f(p_f - p_T)A(r-2), \tag{2}$$

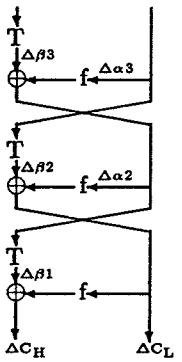
where

$$A(r) := \sum_{\Delta x \neq 0} DP_{LB}\big(r, \Delta P \to (0, \Delta x)\big).$$

Inequality (2) shows that if $p_f \geq p_T$ then upper bounds become smaller than p_f^2 but if $p_f < p_T$ then bounds exceed p_f^2. So we suppose $p_f \geq p_T$ to get a negative answer to Conjecture 3.

Proof

$DP_{LB}(r, \Delta P \to \Delta C)$ is expressed as a followng figure of network and an expression.



$$\begin{aligned} DP_{LB}(r, \Delta P \to \Delta C) \\ = \sum_{\Delta\alpha 2, \Delta\alpha 3, \Delta\beta 1, \Delta\beta 2} & DP_{LB}(r-2, \Delta P \to (\Delta\alpha 2, \Delta\alpha 3)) \\ & dp_T(\Delta\alpha 3 \to \Delta\beta 2) dp_f(\Delta\alpha 2 \to \Delta\beta 2 \oplus \Delta C_L) \\ & dp_T(\Delta\alpha 2 \to \Delta\beta 1) dp_f(\Delta C_L \to \Delta\beta 1 \oplus \Delta C_H). \end{aligned}$$

In the case of $\Delta C_L = 0$, $\Delta\beta 1 = \Delta C_H \neq 0$ and the second to last term of right side of the above equation implies $\Delta\alpha 2 \neq 0$. Thus

$$\begin{aligned} DP_{LB}(r, \Delta P \to \Delta C) &\leq p_T \sum_{\Delta\alpha 2 \neq 0, \Delta\alpha 3, \Delta\beta 2} DP_{LB}(r-2, \Delta P \to (\Delta\alpha 2, \Delta\alpha 3)) dp_T(\Delta\alpha 3 \to \Delta\beta 2) dp_f(\Delta\alpha 2 \to \Delta\beta 2) \\ &\leq p_T p_f \sum_{\Delta\alpha 2 \neq 0, \Delta\alpha 3} DP_{LB}(r-2, \Delta P \to (\Delta\alpha 2, \Delta\alpha 3)) \\ &= p_T p_f - p_T p_f \sum_{\Delta\alpha 3 \neq 0} DP_{LB}(r-2, \Delta P \to (0, \Delta\alpha 3)) = p_T p_f - p_T p_f A(r-2). \end{aligned}$$

Last, in the case of $\Delta C_L \neq 0$, as $dp_f(\Delta C_L \to \Delta\beta 1 \oplus \Delta C_H) \leq p_f$

$$\begin{aligned} & DP_{LB}(r, \Delta P \to \Delta C) \\ \leq\ & p_f \sum_{\Delta\alpha 2, \Delta\alpha 3, \Delta\beta 2} \overbrace{DP_{LB}(r-2, \Delta P \to (\Delta\alpha 2, \Delta\alpha 3)) dp_T(\Delta\alpha 3 \to \Delta\beta 2) dp_f(\Delta\alpha 2 \to \Delta\beta 2 \oplus \Delta C_L)}^{\star\star\star} \\ =\ & p_f \sum_{\Delta\alpha 2 = 0, \Delta\alpha 3, \Delta\beta 2} \star\star\star + p_f \sum_{\Delta\alpha 2 \neq 0, \Delta\alpha 3, \Delta\beta 2} \star\star\star . \end{aligned}$$

When $\Delta\alpha 2 = 0$, as $\Delta\beta 2 = \Delta C_L \neq 0$ and $\Delta\alpha 3 \neq 0$ so

$$p_f \sum_{\Delta\alpha 2 = 0, \Delta\alpha 3, \Delta\beta 2} \star\star\star \leq p_f p_T \sum_{\Delta\alpha 3 \neq 0} DP_{LB}(r-2, \Delta P \to (0, \Delta\alpha 3)) = p_T p_f A(r-2).$$

Moreover as

$$\begin{aligned} p_f \sum_{\Delta\alpha 2 \neq 0, \Delta\alpha 3, \Delta\beta 2} \star\star\star &\leq p_f^2 \sum_{\Delta\alpha 2 \neq 0, \Delta\alpha 3} DP_{LB}(r-2, \Delta P \to (\Delta\alpha 2, \Delta\alpha 3)) \\ &= p_f^2 - p_f^2 A(r-2), \end{aligned}$$

we can obtain

$$DP_{LB}(r, \Delta P \to \Delta C) \leq p_f^2 - p_f^2 A(r-2) + p_f p_T A(r-2). \qquad \text{(Q.E.D.)}$$

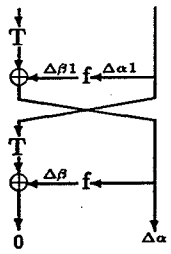
Series $\{A(r)\}$ define a kind of asymptotic inequalities as follows.

Lemma 2

$$\begin{aligned} A(r) &\leq p_f - p_f A(r-1), \\ -A(r) &\leq -\varepsilon_f + \varepsilon_f A(r-1). \end{aligned}$$

Proof

In this cases, $A(r)$ is expressed as a following figure of network and an expression to be evaluated.



$$
\begin{aligned}
A(r) &= \sum_{\Delta\alpha\neq 0,\Delta\alpha 1,\Delta\beta} DP_{LB}(r-1,\Delta P\rightarrow(\Delta\alpha,\Delta\alpha 1)) \\
&\qquad dp_T(\Delta\alpha 1\rightarrow\Delta\beta)dp_f(\Delta\alpha\rightarrow\Delta\beta) \\
&\leq p_f \sum_{\Delta\alpha\neq 0,\Delta\alpha 1} DP_{LB}(r-1,\Delta P\rightarrow(\Delta\alpha,\Delta\alpha 1)) \\
&= p_f - p_f \sum_{\Delta\alpha 1\neq 0} DP_{LB}(r-1,\Delta P\rightarrow(0,\Delta\alpha 1)).
\end{aligned}
$$

In the same way we can get

$$A(r)\geq \varepsilon_f - \varepsilon_f A(r-1). \qquad \text{(Q.E.D.)}$$

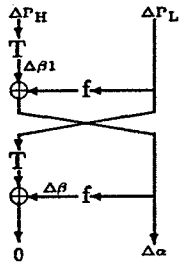
To solve these asymptotic inequalities, the initial values are estimated as;

Lemma 3

$$
\begin{aligned}
A(2) &\leq p_f, \\
-A(2) &\leq -\varepsilon_f + \varepsilon_f p_f.
\end{aligned}
$$

Proof

$A(2)$ is expressed as a following figure of a network and an expression.



$$
\begin{aligned}
A(2) = \sum_{\Delta\alpha\neq 0,\Delta\beta,\Delta\beta 1} & dp_T(\Delta P_H\rightarrow\Delta\beta 1)dp_f(\Delta P_L\rightarrow\Delta\beta 1\oplus\Delta\alpha) \\
& dp_T(\Delta P_L\rightarrow\Delta\beta)dp_f(\Delta\alpha\rightarrow\Delta\beta).
\end{aligned}
$$

In this case as $dp_f(\Delta\alpha\rightarrow\Delta\beta)\leq p_f$ and $-dp_f(\Delta\alpha\rightarrow\Delta\beta)\leq -\varepsilon_f$ so

$$
\begin{aligned}
A(2) &\leq p_f \sum_{\Delta\alpha\neq 0,\Delta\beta 1} dp_T(\Delta P_H\rightarrow\Delta\beta 1)dp_f(\Delta P_L\rightarrow\Delta\beta 1\oplus\Delta\alpha) \\
&= p_f - p_f\sum_{\Delta\beta 1} dp_T(\Delta P_H\rightarrow\Delta\beta 1)dp_f(\Delta P_L\rightarrow\Delta\beta 1) \\
&\leq p_f \qquad (\text{when } \Delta P_L = 0),
\end{aligned}
$$

$$
\begin{aligned}
-A(2) &\leq -\varepsilon_f \sum_{\Delta\alpha\neq 0,\Delta\beta 1} dp_T(\Delta P_H\rightarrow\Delta\beta 1)dp_f(\Delta P_L\rightarrow\Delta\beta 1\oplus\Delta\alpha) \\
&= -\varepsilon_f + \varepsilon_f\sum_{\Delta\beta 1} dp_T(\Delta P_H\rightarrow\Delta\beta 1)dp_f(\Delta P_L\rightarrow\Delta\beta 1) \\
&\leq -\varepsilon_f + \varepsilon_f p_f\sum_{\Delta\beta 1} dp_T(\Delta P_H\rightarrow\Delta\beta 1) \quad (\text{from the assumption of } p_f\geq p_T \text{ and when } \Delta P_L\neq 0) \\
&= -\varepsilon_f + \varepsilon_f p_f.
\end{aligned}
$$

(Q.E.D.)

In case of $-A(2)$ estimation the assumption $p_f\geq p_T$ is essentially used to obtain the result.

From Lemma 2 and Lemma 3 series $\{-A(r)\}$ are solved as

$$
\begin{aligned}
-A(2k) &\leq -(1-p_f)\sum_{i=1}^{k}\varepsilon_f{}^i p_f^{i-1} \quad \text{for any } k\geq 1, \\
-A(2k-1) &\leq -(1-p_f)\sum_{i=1}^{k-1}\varepsilon_f{}^i p_f^{i-1} \quad \text{for any } k\geq 1.
\end{aligned}
$$

As it is easy to confirm that $-A(1) \leq 0$ and $\sum_{i=1}^{0}$ is 0 so the second inequality is true in the case of $k = 1$. From these results the right side of two inequalities of Lemma 1 are expressed from these parameters of p_f, p_T and ε_f, and results are obtained as follows.

$$DP_{LB}(2k-1, \Delta P \to \Delta C)|_{(\Delta C_L=0)} \leq p_T p_f - (1-p_f)p_T \sum_{i=1}^{k-2} {\varepsilon_f}^i p_f^i \quad \text{for any } k \geq 2,$$

$$DP_{LB}(2k, \Delta P \to \Delta C)|_{(\Delta C_L=0)} \leq p_T p_f - (1-p_f)p_T \sum_{i=1}^{k-1} {\varepsilon_f}^i p_f^i \quad \text{for any } k \geq 1.$$

and from the assumption of $p_f \geq p_T$

$$DP_{LB}(2k-1, \Delta P \to \Delta C)|_{(\Delta C_L\neq 0)} \leq p_f^2 - (1-p_f)(p_f - p_T) \sum_{i=1}^{k-2} {\varepsilon_f}^i p_f^i \quad \text{for any } k \geq 2,$$

$$DP_{LB}(2k, \Delta P \to \Delta C)|_{(\Delta C_L\neq 0)} \leq p_f^2 - (1-p_f)(p_f - p_T) \sum_{i=1}^{k-1} {\varepsilon_f}^i p_f^i \quad \text{for any } k \geq 1.$$

Thus we obtain the results of Theorem 7.

4.2 Estimation of Provable Security to Differential Probability of The Other Types

In this subsection the estimation results of differential provability of the other types are mentioned.

4.2.1 B-type

First B-type cipher is examined with using parameter ε_f to show an affirmative answer to Conjecture 1 as follows.

Theorem 8

$$\begin{aligned} DP_B(2k) &\leq \textstyle\sum_{i=2}^{k-1} p_f^i + 2p_f^k - \sum_{i=1}^{k-1} ({p_f}^{i+1} - p_f^{k+1}){\varepsilon_f}^i && \textit{for any } k \geq 1. \\ DP_B(2k-1) &\leq max\{ \textstyle\sum_{i=2}^{k-2} {p_f}^i + 2p^{k-1} - \sum_{i=1}^{k-2} ({p_f}^{i+1} - p_f^k){\varepsilon_f}^i, \\ &\quad \textstyle\sum_{i=2}^{k-1} {p_f}^i + 2p_f^k - \sum_{i=1}^{k-1} ({p_f}^{i+1} - p_f^k){\varepsilon_f}^i\} && \textit{for any } k \geq 2. \end{aligned}$$

In this case, upper bound also decreases as the number of rounds increases but this bound does not become smaller than p_f^2. This is caused by the injectivity of function f. This bound becomes smaller than $2p_f^2$ and the limit of upper bound is $\frac{p_f^2}{1-p_f} - \frac{\varepsilon_f}{1-\varepsilon_f p_f} p_f^2 (> {p_f}^2)$. Gap from $2p_f^2$ is $\frac{1-2p_f}{1-p_f} p_f^2 + \frac{\varepsilon_f}{1-\varepsilon_f p_f} p_f^2$.

4.2.2 RB-type

Theorem 9

$$DP_{RB}(r) \leq p_S p_f \quad \textit{for any } r \geq 2.$$

As same as Theorem 6 estimation of 2-round RB-type could be compared with estimations of 4-round R-type or 4-round B-type and as $p_S p_f \leq max\{p_S^2, 2p_f^2\}$ RB-type seems a more secure Feistel ciphers than R-type and B-type. As same as in case of Theorem 7, by using the parameter ε_f, the examination result of RB-types is given as follows.

Theorem 10

$$\begin{aligned} DP_{RB}(2k) &\leq p_f p_S - p_S(1-p_f) \textstyle\sum_{i=1}^{k-1} {\varepsilon_f}^i {p_f}^i && \textit{for any } k \geq 1. \\ DP_{RB}(2k-1) &\leq p_f p_S - p_S(1-p_f) \textstyle\sum_{i=1}^{k-2} {\varepsilon_f}^i {p_f}^i && \textit{for any } k \geq 2. \end{aligned}$$

The limit of upper bound is $p_f p_S - \frac{1-p_f}{1-\varepsilon_f p_f} \varepsilon_f p_f p_S$ and a gap from the result of Theorem 9 is $\frac{1-p_f}{1-\varepsilon_f p_f} \varepsilon_f p_f p_S$.

4.2.3 LR-type

Theorem 11

$$DP_{LR}(r) \leq max\{{p_S}^2, p_S p_T, {p_T}^2\} \quad \textit{for any } r \geq 3.$$

In this case, because of $max\{p_S^2, p_S p_T, p_T^2\} = max\{p_S^2, p_T^2\}$ the security of LR-type is not so different as that of L-type or R-type and using minimum probability as ε_f is not effective to estimate bound more exactely. The upper bounds are same as when using parameters of ε_S and ε_T.

4.2.4 LRB-type

Theorem 12

$$DP_{LRB}(r) \leq max\{p_f p_S, p_f p_T\} \qquad \textit{for any } r \geq 2.$$

In this case, as $max\{p_f p_S, p_f p_T\} \leq max\{2p_f^2, p_S^2, p_T^2\}$ so the 2-round LRB-type seems more secure Feistel ciphers than 6-round L-type, R-type and B-type because of same reasons as LB-type . With using parameter ε_f, the upper bound is strictly estimated as following.

Theorem 13

$$\begin{aligned}
DP_{LRB}(2k) \leq max\{ & \; p_T p_f - p_T p_f (1-p_f) \textstyle\sum_{i=1}^{k-1} {\varepsilon_f}^i {p_f}^{i-1}, \\
& \; p_S p_f - p_S p_f (1-p_f) \textstyle\sum_{i=1}^{k-1} {\varepsilon_f}^i {p_f}^{i-1}\} \qquad \textit{for any } k \geq 1. \\
DP_{LRB}(2k-1) \leq max\{ & \; p_T p_f - p_T p_f (1-p_f) \textstyle\sum_{i=1}^{k-2} {\varepsilon_f}^i {p_f}^{i-1}, \\
& \; p_S p_f - p_S p_f (1-p_f) \textstyle\sum_{i=1}^{k-1} {\varepsilon_f}^i {p_f}^{i-1}\} \qquad \textit{for any } k \geq 2.
\end{aligned}$$

The limit of upper bound is $max\{\; p_f p_T - \dfrac{1-p_f}{1-\varepsilon_f p_f}\varepsilon_f p_f p_T, p_f p_S - \dfrac{1-p_f}{1-\varepsilon_f p_f}\varepsilon_f p_f p_S\}$ and a gap from Theorem 12 is $\dfrac{1-p_f}{1-\varepsilon_f p_f}\varepsilon_f p_f p_T$ or $\dfrac{1-p_f}{1-\varepsilon_f p_f}\varepsilon_f p_f p_S$.

4.3 Estimation of Provable Security to Linear Probability

As described about the duality structure in Theorem 1, results of estimation of r-round probability of linear hulls are easily obtained as follows.

Theorem 14 *The r-round probability of linear hull on LB-types is estimated as*

$$LP_{LB}(r) \leq q_T q_f \quad \textit{for all } r \geq 2.$$

Moreover when using minimum probability δ_f, estimation is given by

$$\begin{aligned}
LP_{LB}(2k) &\leq q_f q_T - q_T(1-q_f) \textstyle\sum_{i=1}^{k-1} {\delta_f}^i {q_f}^i \qquad \textit{for any } k \geq 1. \\
LP_{LB}(2k-1) &\leq q_f q_T - q_T(1-q_f) \textstyle\sum_{i=1}^{k-2} {\delta_f}^i {q_f}^i \qquad \textit{for any } k \geq 2.
\end{aligned}$$

Theorem 15 *In case of RB-types r-round probability of linear hull is estimated as*

$$LP_{RB}(r) \leq max\{q_S q_f, q_f^2\} \qquad \textit{for any } r \geq 2.$$

and more strictly if $q_f \geq q_S$ then it can be estimated as

$$\begin{aligned}
LP_{RB}(2k) \leq max\{ & \; q_S q_f - (1-q_S) q_S \textstyle\sum_{i=1}^{k-1} {\delta_f}^i {q_f}^i, \\
& \; {q_f}^2 - (1-q_f)(q_f - q_S) \textstyle\sum_{i=1}^{k-1} {\delta_f}^i {q_f}^i\} \qquad \textit{for any } k \geq 1. \\
LP_{RB}(2k-1) \leq max\{ & \; q_S q_f - (1-q_f) q_S \textstyle\sum_{i=1}^{k-2} {\delta_f}^i {q_f}^i, \\
& \; {q_f}^2 - (1-q_f)(q_f - q_S) \textstyle\sum_{i=1}^{k-2} {\delta_f}^i {q_f}^i\} \qquad \textit{for any } k \geq 2.
\end{aligned}$$

Theorem 16

$$LP_{LR}(r) \leq max\{{q_S}^2, q_S q_T, {q_T}^2\}_{196} \textit{ for any } r \geq 3.$$

Table 1 Estimation results

Type	B-type	L-type	R-type	LB-type	RB-type	LR-type	LRB-type
Differential	$2p_f^2$	p_T^2	p_S^2	$max\{p_f p_T, p_f^2\}$	$p_f p_S$	$max\{p_S^2, p_S p_T, p_T^2\}$	$max\{p_f p_S, p_f p_T\}$
Linear hull	$2q_f^2$	q_T^2	q_S^2	$q_f q_T$	$max\{q_f q_S, q_f^2\}$	$max\{q_S^2, q_S q_T, q_T^2\}$	$max\{q_f q_S, q_f q_T\}$
Round dependency	detected	nondetected	nondetected	detected	detected	nondetected	detected

Theorem 17

$$LP_{LRB}(r) \leq max\{q_f q_S, q_f q_T\} \qquad \textit{for any } r \geq 2.$$

By more strict estimation with using δ_f estimation becomes

$$\begin{aligned} LP_{LRB}(2k) \leq max\{ & q_S q_f - q_S q_f (1-q_f) \textstyle\sum_{i=1}^{k-1} {\delta_f}^i {q_f}^{i-1}, \\ & q_T q_f - q_T q_f (1-q_f) \textstyle\sum_{i=1}^{k-1} {\delta_f}^i {q_f}^{i-1}\} \qquad \textit{for any } k \geq 1. \\ LP_{LRB}(2k-1) \leq max\{ & q_S q_f - q_S q_f (1-q_f) \textstyle\sum_{i=1}^{k-2} {\delta_f}^i {q_f}^{i-1}, \\ & q_T q_f - q_T q_f (1-q_f) \textstyle\sum_{i=1}^{k-1} {\delta_f}^i {q_f}^{i-1}\} \qquad \textit{for any } k \geq 2. \end{aligned}$$

Acknowledgements

We would like to thank S.Tsujii(leader of the Information & Communication Project) for his well advice and T.Saiki(staff of Telecommunications Advancement Organization of Japan) for his support to our activity.

5 Conclusion

We have introduced Generalized Feistel Ciphers and classified them to estimate the upper bounds to the probabilities of r-round differentials and r-round linear hulls. Table 1 shows estimation results.

Estimation duality between the differential probabilities and the linear hull probabilities is found in this table. In these cases of LB-type and RB-type upper bounds become p_f^2 or q_f^2 without supposing injectiveness of a function f. We have also shown that if Generalized Feistel Ciphers include a B-type random function then provable security depends on the round number of these block ciphers. Each estimation results of upper bounds are expressed as polynomials of p_* or q_* and the lowest degree of these polynomial seems to be less than or equal to 2 and it may be true for any types of Feistel ciphers, each round of which has 2-inputs and 2-outputs. Moreover without describing in this paper, by using mixed random functions in Generalized Feistel Ciphers as differential and linear approximation equation per one-round are supposed to be more difficult to solve than when using only one random function so Generalized Feistel Ciphers may be useful in strengthening these usual block ciphers.

References

[AO97] K.Aoki and K. Ohta. *Strict evaluation of the maximum average of differential probability and the maximum average of linear probability* IEICE Trans. Volume E80-A, Number 1, pp.2–8. (1997).

[BS91] E.Biham and A. Shamir. *Differential Cryptanalysis of DES-like Cryptosystems.* Journal of Cryptology, Volume 4, Number 1, pp.3–72, Springer International, 1991.

[CV94] F.Chabaud and S.Vaudenay. *Links Between Differential and Linear Cryptanalysis.* Advances in Cryptology–EUROCRYPT'94, Lecture Notes in Computer Sciences 950, pp.356–365, Springer-Verlag, 1995.

[F73] H. Feistel. *Cryptography and computer privacy.* in Scientific American, Vol.228, pp.15-23 (1973).

[FNS75] H. Feistel, W.A. Notz and J.L. Smith. *Some cryptographic techniques for machine-to-machine data communications.* in Proceedings of IEEE, Vol.63, No. 11, pp.1545-1554 (1975).

[LMM91] X.Lai, J.L.Massey and S.Murphy. *Markov Ciphers and Differential Cryptanalysis.* Advances in Cryptology–EUROCRYPT'91, Lecture Notes in Computer Sciences 547, pp.17–38, Springer-Verlag, 1992.

[Mat93] M.Matsui. *Linear Cryptanalysis Method for DES Cipher.* Advances in Cryptology–EUROCRYPT'93, Lecture Notes in Computer Sciences 765, pp.386–397, Springer-Verlag, 1994.

[Mat94] M.Matsui. *On Correlation Between the Order of S-boxes and the Strength of DES.* Advances in Cryptology–EUROCRYPT'94, Lecture Notes in Computer Sciences 950, pp.366–375, Springer-Verlag, 1995.

[Mat96] M.Matsui. *New Structure of Block Ciphers with Provable Security against Differential and Linear Cryptanalysis.* In Proceedings of the third international workshop of fast software encryption, Lecture Notes in Computer Science 1039, pp.205–218, Springer-Verlag, 1996.

[NBS77] National Bureau of Standards, NBS FIPS PUB 46, *Data Encryption Standard* , U.S.Department of Commerce (Jan. 1977).

[Nyb94] K.Nyberg. *Linear Approximation of Block Ciphers.* Advances in Cryptology–EUROCRYPT'94, Lecture Notes in Computer Sciences 950, pp.439–444, Springer-Verlag, 1995.

[NK95] K.Nyberg and L.R.Knudsen. *Provable Security Against a Differential Attack.* Journal of Cryptology, Volume 8, Number 1, pp.27–37, Springer International, 1995.

[Nyb96] K.Nyberg. *Generalized Feistel Networks.* Advances in Cryptology–CRYPTO'96, Lecture Notes in Computer Sciences 1163, pp.91–104, Springer-Verlag, 1996.

[OG94] L.O'Connor and J.D.Golic. *A Unified Markov Approach to Differential and Linear Cryptanalysis.* Advances in Cryptology–ASIACRYPT'94, Lecture Notes in Computer Sciences 917, pp.387–397, Springer-Verlag, 1995.

[OMA95] K.Ohta, S.Moriai and K.Aoki. *Improving the Search Algorithm for the Best Linear Expression.* Advances in Cryptology–CRYPTO'95, Lecture Notes in Computer Science 963, pp.157–170, Springer-Verlag, 1995.

[S-B83] I.S.-Bichl, *On the design and analysis of new cipher systems related to the DES*, Tech. Rept. Linz Univ. 1993.

[Sch93] B. Schneiner, *Description of a New Variable-Length Key, 64-bit Block Chiper (Blowfish).* Fast Software Encryption, Springer-Verlag(1993).

Appendix A

To construct a Key scheduler which generates m_1-bit round key k_1, m_2-bit round key k_2 and m_3-bit round key k_3 from M-bit secret key, we modified the well known key scheduler structure of DES as following.

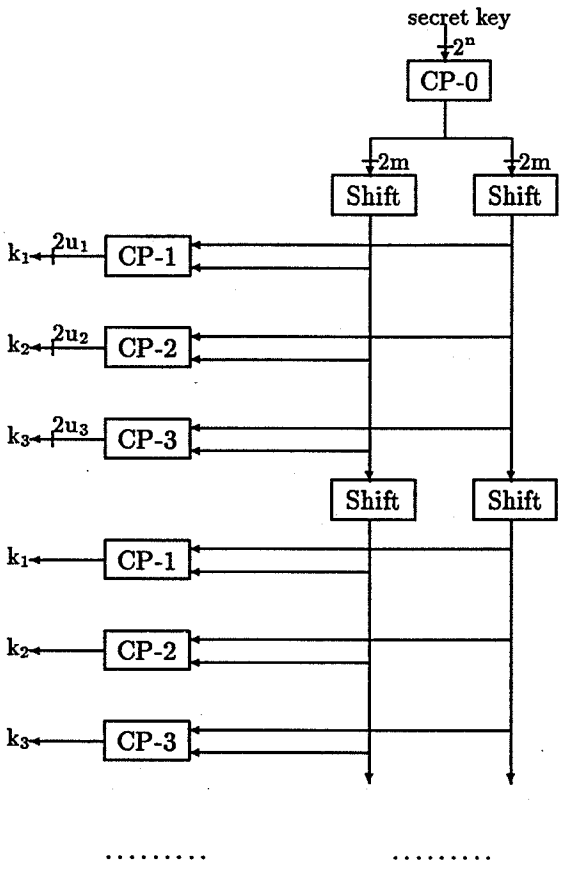


Figure: Structure of Generalized Key scheduler

In this figure, in spite of only using one compression permutation CP-3 which generates $m_3(= 2u_3)$-bit round key k_3, two compression permutations, CP-1 which generates $m_1(= 2u_1)$-bit round key k_1 and CP-2 which generates $m_2(= 2u_2)$-bit round key k_2, are added.

A bit of the story of PGP, future directions, and some public policy thoughts on crypto

Phil Zimmermann
PGP Inc., San Mateo, CA

July 28, 1997

Abstract

An informal talk on the history of PGP from Zimmermann's personal perspective, with some discussion of the relevent public policy issues, and some technical information of the new PGP architecture.