Markov Ciphers and Differential Cryptanalysis

Xuejia Lai James L. Massey

Institute for Signal and Information Processing
Swiss Federal Institute of Technology
CH–8092 Zürich, Switzerland

Sean Murphy
Royal Holloway & Bedford New College
University of London, Egham, Surrey TW20 0EX

Abstract

This paper considers the security of iterated block ciphers against the differential cryptanalysis introduced by Biham and Shamir. Differential cryptanalysis is a chosen-plaintext attack on secret-key block ciphers that are based on iterating a cryptographically weak function r times (e.g., the 16-round Data Encryption Standard (DES)). It is shown that the success of such attacks on an r-round cipher depends on the existence of (r-1)-round differentials that have high probabilities, where an i-round differential is defined as a couple (α, β) such that a pair of distinct plaintexts with difference α can result in a pair of i-th round outputs that have difference β, for an appropriate notion of "difference". The probabilities of such differentials can be used to determine a lower bound on the complexity of a differential cryptanalysis attack and to show when an r-round cipher is not vulnerable to such attacks. The concept of "Markov ciphers" is introduced for iterated ciphers because of its significance in differential cryptanalysis. If an iterated cipher is Markov and its round subkeys are independent, then the sequence of differences at each round output forms a Markov chain. It follows from a result of Biham and Shamir that DES is a Markov cipher. It is shown that, for the appropriate notion of "difference", the Proposed Encryption Standard (PES) of Lai and Massey, which is an 8-round iterated cipher, is a Markov cipher, as are also the mini-version of PES with block length 8, 16 and 32 bits. It is shown that PES(8) and PES(16) are immune to differential cryptanalysis after sufficiently many rounds. A detailed cryptanalysis of the full-size PES is given and shows that the very plausibly most probable 7-round differential has a probability about 2^{-58}. A differential cryptanalysis attack of PES(64) based on this differential is shown to

D.W. Davies (Ed.): Advances in Cryptology - EUROCRYPT '91, LNCS 547, pp. 17-38, 1991.

require all 2^{64} possible encryptions. This cryptanalysis of PES suggested a new design principle for Markov ciphers, viz., that their transition probability matrices should not be symmetric. A minor modification of PES, consistent with all the original design principles, is proposed that satisfies this new design criterion. This modified cipher, called Improved PES (IPES), is described and shown to be highly resistant to differential cryptanalysis.

1. Introduction

Many secret-key block ciphers are cryptosystems based on iterating a cryptographically weak function several times. Each iteration is called a round. The output of each round is a function of the output of the previous round and of a subkey derived from the full secret key by a key-schedule algorithm. Such a secret-key block cipher with r-iterations is called an r-round iterated cipher. For example, the well-known Data Encryption Standard (DES) is a 16-round iterated cipher.

Differential cryptanalysis, introduced by Biham and Shamir in [1], is a chosen-plaintext attack to find the secret key of an iterated ciphers. It analyzes the effect of the "difference" of a pair of plaintexts on the "difference" of succeeding round outputs in an r-round iterated cipher. In Section 2, we describe differential cryptanalysis of a general r-round iterated cipher in terms of (r-1)-round "differentials" instead of in terms of the "i-round characteristics" used in [1]. The hypothesis of stochastic equivalence, which has been implicitly assumed in differential cryptanalysis, is explicitly formulated in Section 2. It is pointed out that one of the two prerequisites for differential cryptanalysis to succeed on an r-round cipher is the existence of an (r-1)-round differential with high probability, and it is shown that a lower bound on the complexity of differential cryptanalysis can be obtained from the maximum differential probability.

In Section 3, Markov ciphers are defined as iterated ciphers whose round functions satisfy the condition that the differential probability is independent of the choice of one of the component plaintexts under an appropriate definition of difference. It is shown that, for a Markov cipher with independent subkeys, the sequence of round differences forms a Markov chain. It follows from a result of Biham and Shamir [1] that DES is a Markov cipher. The study of differential cryptanalysis for an r-round Markov cipher is reduced to the study of the transition probabilities created by its round function. In particular, Markov chain techniques can be used to show whether the cipher is secure against differential cryptanalysis after sufficiently many rounds.

At Eurocrypt'90, a new iterated cipher, the Proposed Encryption Standard (PES) was introduced by Lai and Massey [2]. The PES contains 8 rounds plus an output transformation. In Section 4, standard PES with block length 64 bits and mini-versions of PES with block length 8, 16 and 32 are considered. These are all shown to be Markov ciphers. The ciphers PES(8) and PES(16) are shown to be immune to differential cryptanalysis after sufficiently many rounds. A detailed cryptanalysis of PES(64), given in the Appendix, shows that the very plausibly most likely one-round differential has probability

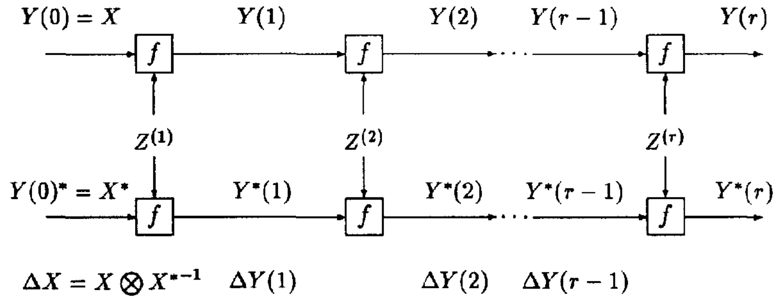


Figure 1: Encrypting a pair of plaintexts with an r-round iterated cipher

about 2^{-9}, which leads to a 7-round differential with probability about 2^{-58}. Differential cryptanalysis of PES(64) based on this differential requires the cryptanalyst to perform all 2^{64} possible encryptions. The attacker thus obtains the secret key after 2^{64} encryptions, which is much less than the 2^{128} encryptions of an exhaustive key search; however, the 2^{64} encryptions specify the entire mapping from plaintext to ciphertext determined by the secret key and hence the attacker has no need to find the actual secret key.

The cryptanalysis of PES shows that the symmetry of transition probability matrix of its Markov chain is responsible for the "undesirably large" probability of its most probable 7-round differential. This suggests a new design principle for Markov ciphers, viz., that their transition probability matrices should not be symmetric. A minor modification of PES, called Improved PES (IPES), was suggested by this new design principle and is described in Section 5. It is shown that this modification substantially improves the security of PES without violating any of the earlier design principles used for PES.

2. Differential Cryptanalysis of Iterated Ciphers

Throughout this paper, we consider the encryption of a pair of *distinct* plaintexts by an r-round iterated cipher as shown schematically in Fig.1. In this figure, the round function $Y = f(X, Z)$ is such that, for every round subkey Z, $f(\cdot, Z)$ establishes a one-to-one correspondence between the round input X and the round output Y. Let the "difference" ΔX between two plaintexts (or two ciphertexts) X and X^* be defined as

$$\Delta X = X \otimes X^{*-1},$$

where $\otimes$ denotes a specified group operation on the set of plaintexts (= set of ciphertexts) and X^{*-1} denotes the inverse of the element X^* in the group. The round function $Y = f(X, Z)$ is said to be *cryptographically weak* if, given a few triples $(\Delta X, Y, Y^*)$, it is feasible (in most cases) to determine the subkey Z.

From the pair of encryptions, one obtains the sequence of differences $\Delta Y(0)$, $\Delta Y(1), \ldots,$ $\Delta Y(r)$ where $Y(0) = X$ and $Y^*(0) = X^*$ denote the plaintext pair [so that $\Delta Y(0) = \Delta X$]

and where $Y(i)$ and $Y^*(i)$ for $(0 < i < r)$ are the outputs of the i-th round, which are also the inputs to the (i+1)-th round. The subkey for the i-th round is denoted as $Z^{(i)}$. In the following discussion, we always assume that $X \neq X^*$ because, when $X = X^*$, all $\Delta Y(i)$ would equal the *neutral* element e of the group, which case is of no interest for differential cryptanalysis.

Differential cryptanalysis exploits the fact that the round function f in an iterated cipher is usually cryptographically weak. Thus, if the ciphertext pair is known and the difference of the pair of inputs to the last round can somehow be obtained, then it is possible to determine (some substantial part of) the subkey of the last round. In differential cryptanalysis, this is achieved by *choosing* plaintext pairs (X, X^*) with a specified difference α such that the difference $\Delta Y(r-1)$ of the pair of inputs to the last round will take on a particular value β with high probability. Based on this idea, we make the following definition.

Definition. An *i-round differential* is a couple (α, β), where α is the difference of a pair of distinct plaintexts X and X^* and where β is a possible difference for the resulting i-th round outputs $Y(i)$ and $Y^*(i)$. The *probability of an i-round differential* (α, β) is the conditional probability that β is the difference $\Delta Y(i)$ of the ciphertext pair after i rounds given that the plaintext pair (X, X^*) has difference $\Delta X = \alpha$ *when the plaintext* X *and the subkeys* $Z^{(1)}, \ldots, Z^{(i)}$ *are independent and uniformly random.* We denote this differential probability by $P(\Delta Y(i) = \beta | \Delta X = \alpha)$.

The basic procedure of a differential cryptanalysis *attack* on an r-round iterated cipher can be summarized as follows:

1) Find an (r-1)-round differential (α, β) such that $P(\Delta Y(r-1) = \beta | \Delta X = \alpha)$ has maximum, or nearly maximum, probability.

2) Choose a plaintext X uniformly at random and compute X^* so that the difference ΔX between X and X^* is α. Submit X and X^* for encryption under the actual key Z. From the resultant ciphertexts $Y(r)$ and $Y^*(r)$, find every possible value (if any) of the subkey $Z^{(r)}$ of the last round corresponding to the anticipated difference $\Delta Y(r-1) = \beta$. Add one to the count of the number of appearances of each such value of the subkey $Z^{(r)}$.

3) Repeat 2) until one or more values of the subkey $Z^{(r)}$ are counted significantly more often than the others. Take this most-often-counted subkey, or this small set of such subkeys, as the cryptanalyst's decision for the actual subkey $Z^{(r)}$.

Note that, in a differential cryptanalysis attack, all the subkeys are *fixed* and only the plaintext can be randomly chosen. In the computation of a differential probability, however, the plaintext and all subkeys are independent and uniformly random. In preparing a differential cryptanalysis attack, one uses the computed differential probabilities to determine which differential to use in the attack; hence, one is tacitly making the following hypothesis.

Hypothesis of Stochastic Equivalence. *For an (r-1)-round differential* (α,β),

$$P(\Delta Y(r-1)=\beta|\Delta X=\alpha)\approx P(\Delta Y(r-1)=\beta|\Delta X=\alpha, Z^{(1)}=\omega_1,...,Z^{(r-1)}=\omega_{r-1})$$

for almost all subkey values $(\omega_1,...,\omega_{r-1})$.

From the above description of a differential cryptanalysis attack and from the fact that there are 2^m-1 possible values of $\Delta Y(r-1)$, one deduces the following result.

Suppose the hypothesis of stochastic equivalence is true, then an r-round cipher with independent subkeys is vulnerable to differential cryptanalysis if and only if the round function is weak and there exists an (r-1)-round differential (α,β) *such that* $P(\Delta Y(r-1)=\beta|\Delta X=\alpha)\gg 2^{-m}$, *where* m *is the block length of the cipher.*

Let $Comp(r)$ denote the *complexity* of differential cryptanalysis of an r-round cipher which, following [1], is defined as the number of encryptions used.

Theorem 1. (Lower bound on the complexity of a differential cryptanalysis attack on an r-round iterated cipher.)

Suppose the hypothesis of stochastic equivalence is true, then, in an attack by differential cryptanalysis,

$$Comp(r)\geq 2/(p_{max}-\frac{1}{2^m-1}) \text{ where } p_{max}=\max_{\alpha}\max_{\beta}P(\Delta Y(r-1)=\beta|\Delta X=\alpha),$$

and where m *is the block length of the plaintext In particular, if* $p_{max}\approx\frac{1}{2^m-1}$, *then a differential cryptanalysis attack will not succeed.*

Proof. Note that the anticipated value β of the difference $\Delta Y(r-1)$ must certainly be taken on at least once more on the average than a randomly chosen value β' if differential cryptanalysis is to succeed. Thus, $Tp_{max}\geq\frac{T}{2^m-1}+1$ is a necessary condition for success in T trials, where each trial consists in choosing a pair of plaintexts with the specified difference α.

Remark. In [1], differential cryptanalysis of DES was described in terms of "i-round characteristics". In our notation, an i-round characteristic as defined in [1] is an $(i+1)$-tuple $(\alpha,\beta_1,...,\beta_i)$ considered as a possible value of $(\Delta X,\Delta Y(1),...,\Delta Y(i))$. Thus, a one-round characteristic coincides with a one-round differential and an i-round characteristic determines a sequence of i differentials, $(\Delta X,\Delta Y(j))=(\alpha,\beta_j)$. The probability of an i-round characteristic is defined in [1] as

$$P(\Delta Y(1)=\beta_1,\Delta Y(2)=\beta_2,..,\Delta Y(i)=\beta_i|\Delta X=\alpha)$$

where the plaintext X and the subkeys $Z^{(1)},...,Z^{(i)}$ are independent and uniformly random. We use the notion of differentials rather than characteristics because, in the differential cryptanalysis of an r-round cipher, only the knowledge of $\Delta Y(r-1)$ is required for determining the subkey $Z^{(r)}$, no matter what the intermediate differences

$\Delta Y(j), 1 \leq j < r-1$, may be. The most probable differential will have in general a probability greater than that of the most probable characteristic (although for DES with a small number of rounds, the two probabilities are roughly the same). Thus, by using differential probabilities rather than characteristic probabilities, we consider in fact the true probability that differential cryptanalysis will succeed, not just a lower bound on this probability. This is why we were able to derive a *lower bound* on the complexity of a differential cryptanalysis attack from the probability of differentials.

3. Markov Ciphers

In this section, a class of iterated ciphers that are especially interesting for differential cryptanalysis will be considered. For such a cipher, the sequence $\Delta Y(0)$, $\Delta Y(1), \ldots, \Delta Y(r)$ forms a Markov chain. Recall that a sequence of discrete random variables $v_0, v_1, \ldots, v_r$ is a *Markov chain* if, for $0 \leq i < r$ (where $r = \infty$ is allowed),

$$P(v_{i+1} = \beta_{i+1} | v_i = \beta_i, v_{i-1} = \beta_{i-1}, \ldots, v_0 = \beta_0) = P(v_{i+1} = \beta_{i+1} | v_i = \beta_i).$$

A Markov chain is called *homogeneous* if $P(v_{i+1} = \beta | v_i = \alpha)$ is independent of i for all α and β. [In what follows, we always assume that the plaintext X is independent of the subkeys $Z^{(1)}, \ldots, Z^{(r)}$.]

Definition. An iterated cipher with round function $Y = f(X, Z)$ is a *Markov cipher* if there is a group operation $\otimes$ for defining differences such that, for all choices of α ($\alpha \neq e$) and β ($\beta \neq e$),

$$P(\Delta Y = \beta | \Delta X = \alpha, X = \gamma)$$

is independent of γ when the subkey Z is uniformly random, or, equivalently, if

$$P(\Delta Y = \beta | \Delta X = \alpha, X = \gamma) = P(\Delta Y(1) = \beta_1 | \Delta X = \alpha)$$

for all choices of γ when the subkey Z is uniformly random.

The following crucial theorem explains the terminology "Markov cipher".

Theorem 2. *If an r-round iterated cipher is a Markov cipher and the r round keys are independent and uniformly random, then the sequence of differences* $\Delta X = \Delta Y(0)$, $\Delta Y(1), \ldots, \Delta Y(r)$ *is a homogeneous Markov chain. Moreover, this Markov chain is stationary if* ΔX *is uniformly distributed over the non-neutral elements of the group.*

Proof. To show that the sequence ΔX, $\Delta Y(1), \ldots, \Delta Y(r)$ is a Markov chain, it is sufficient to show for the second round that

$$P(\Delta Y(2) = \beta_2 | \Delta Y(1) = \beta_1, \Delta X = \alpha) = P(\Delta Y(2) = \beta_2 | \Delta Y(1) = \beta_1).$$

To show this, we note that

$$
\begin{aligned}
&= \sum_{\gamma} P(Y(1) = \gamma|\Delta Y(1) = \beta_1, \Delta X = \alpha)P(\Delta Y(2) = \beta_2|\Delta Y(1) = \beta_1, Y(1) = \gamma, \Delta X = \alpha) \\
&= \sum_{\gamma} P(Y(1) = \gamma|\Delta Y(1) = \beta_1, \Delta X = \alpha)P(\Delta Y(2) = \beta_2|\Delta Y(1) = \beta_1, Y(1) = \gamma) \\
&= \sum_{\gamma} P(Y(1) = \gamma|\Delta Y(1) = \beta_1, \Delta X = \alpha)P(\Delta Y(2) = \beta_2|\Delta Y(1) = \beta_1) \\
&= P(\Delta Y(2) = \beta_2|\Delta Y(1) = \beta_1),
\end{aligned}
$$

where the third equality comes from the fact that $Y(1)$ and $\Delta Y(1)$ together determine both $Y(1)$ and $Y(1)^*$ so that $\Delta Y(2)$ has no further dependence on ΔX when $Y(1)$ and $\Delta Y(1)$ are specified. Because the same round function is used in each round, this Markov chain is homogeneous. For any key $Z = z$, the round function $f(\cdot, z)$ is a bijective mapping from the set of plaintexts to the set of ciphertexts. This bijection induces a bijection from pairs of distinct plaintexts (X, X^*) to pairs of distinct ciphertexts $(Y, Y^*) = (f(X, z), f(X^*, z))$. The fact that X and $\Delta X (\neq e)$ are independent and uni-

Example 1. *DES is a Markov cipher under the definition of difference as $\Delta X = X \bigoplus X^*$ where $\bigoplus$ denotes bitwise XOR. (This is just a restatement of Lemma 1 in [1].)*

For a Markov cipher with independent and uniformly random round subkeys, the probability of an r-round characteristic is given by the Chapman-Kolmogorov equation for a Markov chain as

$$P(\Delta Y(1) = \beta_1, \Delta Y(2) = \beta_2, .., \Delta Y(r) = \beta_r|\Delta X = \beta_0) = \prod_{i=1}^{r} P(\Delta Y(1) = \beta_i|\Delta X = \beta_{i-1}).$$

It follows that the probability of an r-round differential (β_0, β_r) is

$$P(\Delta Y(r) = \beta_r|\Delta X = \beta_0) = \sum_{\beta_1}\sum_{\beta_2}\cdots\sum_{\beta_{r-1}}\prod_{i=1}^{r} P(\Delta Y(1) = \beta_i|\Delta X = \beta_{i-1})$$

where the sums are over all possible values of differences between distinct elements, i.e., over all group elements excepting the neutral element e.

For any Markov cipher, let Π denote the *transition probability matrix* of the homogeneous Markov chain $\Delta X = \Delta Y(0)$, $\Delta Y(1), \ldots, \Delta Y(r)$. The (i, j) entry in Π is

$P(\Delta Y(1) = \alpha_j | \Delta X = \alpha_i)$ where $\alpha_1, \alpha_2, ..., \alpha_M$ is some agreed-upon ordering of the M possible values of ΔX and $M = 2^m - 1$ for an m-bit cipher. Then, for any r, the (i,j) entry in Π^r, $p_{ij}^{(r)}$, equals $P(\Delta Y(r) = \alpha_j | \Delta X = \alpha_i)$, i.e., $p_{ij}^{(r)}$ is just the probability of the r-round differential (α_i, α_j).

The security of iterated cryptosystems is based on the belief that a cryptographically "strong" function can be obtained by iterating a cryptographically "weak" function enough times. For Markov ciphers, one has the following fact.

Theorem 3. *For a Markov cipher of block length m with independent and uniformly random round subkeys, if the semi-infinite Markov chain $\Delta X = \Delta Y(0), \Delta Y(1), ...$ has a "steady-state probability" distribution, i.e., if there is a probability vector $(p_1, p_2, .., p_M)$, such that, for all α_i, $\lim_{r \to \infty} P(\Delta Y(r) = \alpha_j | \Delta X = \alpha_i) = p_j$, then this steady state distribution must be the uniform distribution $(1/M, 1/M, ..., 1/M)$, i.e., $\lim_{r \to \infty} P(\Delta Y(r) = \beta | \Delta X = \alpha) = \frac{1}{2^m - 1}$ for every differential (α, β), so that every differential will be roughly equally likely after sufficiently many rounds. If we assume additionally that the hypothesis of stochastic equivalence holds for this Markov cipher, then, for almost all subkeys, this cipher is secure against a differential cryptanalysis attack after sufficiently many rounds.*

Proof. The theorem follows from the facts that the existence of a steady-state probability distribution implies that a homogeneous Markov chain has a unique stationary distribution, which is the steady-state distribution, and that, according to Theorem 2, the uniform distribution is a stationary distribution.

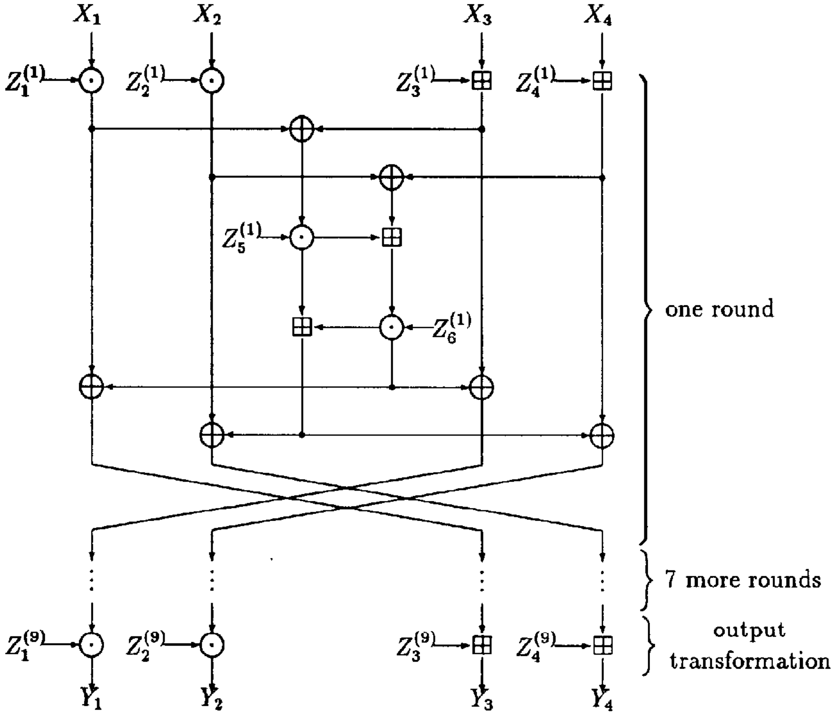
4. Analysis of the block cipher PES

The block cipher PES, proposed by Lai and Massey in [2], is an iterated block cipher based on three group operations on 16-bit subblocks, namely, bitwise-XOR, denoted as $\oplus$; addition modulo 2^{16} of integers represented by 16-bit subblocks, denoted as $\boxplus$; and multiplication modulo $2^{16}+1$ (with the all-zero 16-bit subblock considered as representing 2^{16}), denoted as $\odot$. The encryption process of PES is shown in Fig.2. In order to consider differential cryptanalysis of PES, we must first define "difference".

The encryption of a plaintext pair by an r-round PES can be described as shown in Fig. 3, where X_i and $Z_j^{(i)}$ denote 16-bit subblocks, where $X = (X_1, X_2, X_3, X_4)$, where $Z_A^{(i)} = (Z_1^{(i)}, Z_2^{(i)}, Z_3^{(i)}, Z_4^{(i)})$, where $Z_B^{(i)} = (Z_5^{(i)}, Z_6^{(i)})$, and where we introduce an operation $\otimes$ defined on 64-bit blocks by

$$X \otimes Z_A^{(i)} = (X_1 \odot Z_1^{(i)}, X_2 \odot Z_2^{(i)}, X_3 \boxplus Z_3^{(i)}, X_4 \boxplus Z_4^{(i)}). \qquad (1)$$

Under the operation $\otimes$, the set of all 64-bit blocks forms a group. Let X^{-1} be the inverse of X in this group. Then, for PES, we define the *difference* of two distinct 64-bit blocks X and X^* as $\Delta X = X \otimes X^{*-1}$. The appropriateness of this definition stems from the following fact:



X_i : 16-bit plaintext subblock
Y_i : 16-bit ciphertext subblock
$Z_i^{(r)}$: 16-bit key subblock
$\oplus$: bit-by-bit exclusive-OR of 16-bit subblocks
$\boxplus$: addition modulo 2^{16} of 16-bit integers
$\odot$: multiplication modulo $2^{16} + 1$ of 16-bit integers
with the zero subblock corresponding to 2^{16}

Figure 2: The encryption process of the block cipher PES.

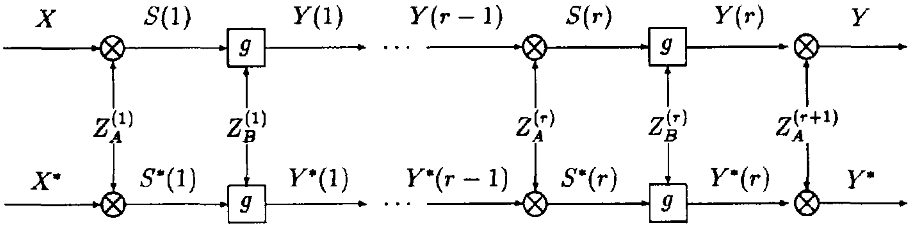


Figure 3: Encrypting a pair of plaintexts with an r-round PES

Lemma 1. *PES is Markov cipher under the definition of difference as* $\Delta X = X \bigotimes X^{*-1}$, *where* $\bigotimes$ *is the group operation defined in (1).*

Proof.

$$
\begin{aligned}
&P(\Delta Y(1) = \beta_1 | \Delta X = \beta_0, X = \gamma) \\
&= \quad P(\Delta Y(1) = \beta_1 | \Delta S(1) = \beta_0, X = \gamma) \\
&= \quad \sum_{\lambda} P(\Delta Y(1) = \beta_1, S(1) = \lambda) | \Delta S(1) = \beta_0, X = \gamma) \\
&= \quad \sum_{\lambda} P(\Delta Y(1) = \beta_1 | \Delta S(1) = \beta_0, X = \gamma, S(1) = \lambda) P(S(1) = \lambda | \Delta S(1) = \beta_0, X = \gamma) \\
&= \quad \sum_{\lambda} P(\Delta Y(1) = \beta_1 | \Delta S(1) = \beta_0, S(1) = \lambda) P(Z_A^{(1)} = \lambda \otimes \gamma^{-1}) \\
&= \quad 2^{-64} \sum_{\lambda} P(\Delta Y(1) = \beta_1 | \Delta S(1) = \beta_0, S(1) = \lambda),
\end{aligned}
$$

which is independent of γ, where we have used the facts that $\Delta S(1) = \Delta X$ since $S(1) = X \otimes Z_A^{(1)}$ and that

$$
P(S(1) = \lambda | \Delta S(1) = \beta_0, X = \gamma) = P(S(1) = \lambda | X = \gamma) = P(Z_A^{(1)} = \lambda \otimes \gamma^{-1}).
$$

The regular structure of PES makes it possible, and insightful, to consider "mini" PES ciphers with shorter block length. A mini PES has the same computational graph as the standard PES shown in Fig.2, but the subblocks are only n bits long (n=2, 4 or 8) rather than 16, and the operations $\bigoplus$,$\boxplus$ and $\odot$ are then the corresponding bitwise XOR, addition modulo 2^n, and multiplication modulo $2^n + 1$. Note that for n=2, 4 and 8, these three operations are still group operations. Thus, the resulting mini PES is a Markov cipher with block length $4n$, by the same argument as for PES(64).

We have been able to prove that, for PES(8) and PES(16), the uniform distribution is indeed the steady-state probability of the sequence of differences. Thus, PES(8) and

PES(16) with sufficiently many rounds are guaranteed secure against differential cryptanalysis. We conjecture that PES(32) and PES(64) also have the uniform distribution as the steady-state probability distribution for their sequences of differences.

In order to find the one-round differential with highest probability, an exhaustive search was performed for the mini ciphers PES(8) and PES(16). The most likely one-round differentials $(\Delta X, \Delta Y(1))$ for PES(8) and PES(16) are:

$$\Delta Y(1) = \Delta X = (\Delta X_1, \Delta X_2, \Delta X_3, \Delta X_4) = (0, 1, odd, 0), \qquad odd \in \{1, 3, .., 2^n - 1\},$$

and each has a probability approximately $2^{-(2n-2)}$. The i-round differentials $(\Delta X, \Delta Y(i))$ that take on these same values also have the greatest probabilities for small $i > 1$. The probabilities of the above i-round differentials for PES(8) and PES(16) are shown in Table 1.

p_{df}	8-bit	16-bit
1-round	1.25×2^{-2}	1.13×2^{-6}
2-round	1.62×2^{-3}	1.47×2^{-10}
3-round	1.07×2^{-3}	1.03×2^{-13}†
4-round	1.43×2^{-4}	1.6×2^{-16}†
5-round	0.97×2^{-4}	*†

Table 1: The probabilities of the (estimated or proved) most probable i-round differentials for PES(8) and PES(16). (* : statistically indistinguishable from 2^{-4n}; † : estimated by statistical test.)

For PES(64), a detailed cryptanalysis (see the Appendix) strongly suggests that the most probable one-round differentials correspond to eight values of ΔX, namely $\alpha_i = (0, 0, 0, \gamma_i)$ (i=1,2,...,8) where $\gamma_1 = 2^{16} - 1$, $\gamma_2 = 1$, $\gamma_3 = 2^{16} - 3$, $\gamma_4 = 3$, $\gamma_5 = 2^{16} - 5$, $\gamma_6 = 5$, $\gamma_7 = 2^{16} - 7$, and $\gamma_8 = 7$. The 8×8 submatrix of the transition probability matrix Π corresponding to these values is shown in the Appendix to be well-approximated by

$$T = 10^{-7} \begin{pmatrix} 0 & 25460 & 12556 & 0 & 0 & 9417 & 698 & 0 \\ 25460 & 0 & 0 & 12556 & 9417 & 0 & 0 & 698 \\ 12556 & 0 & 0 & 0 & 6278 & 0 & 0 & 3139 \\ 0 & 12556 & 0 & 0 & 0 & 6278 & 3139 & 0 \\ 0 & 9417 & 6278 & 0 & 0 & 0 & 0 & 0 \\ 9417 & 0 & 0 & 6278 & 0 & 0 & 0 & 0 \\ 698 & 0 & 0 & 3139 & 0 & 0 & 0 & 0 \\ 0 & 698 & 3139 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Note that the (i, j) entry in T^k is just

$$P(\Delta Y(k) = \alpha_j, \Delta Y(k-1) \in A, \cdots, \Delta Y(1) \in A, |\Delta X = \alpha_i)$$

where $A = \{\alpha_1, \alpha_2, ..., \alpha_8\}$, which is a lower bound on the (i,j) entry of Π^k. One obtains

$$T^7 = 2^{-58} \begin{pmatrix} 0 & 1.22 & 0.53 & 0 & 0 & 0.43 & 0.07 & 0 \\ 1.22 & 0 & 0 & 0.53 & 0.43 & 0 & 0 & 0.07 \\ 0.53 & 0 & 0 & 0.23 & 0.19 & 0 & 0 & 0.03 \\ 0 & 0.53 & 0.23 & 0 & 0 & 0.19 & 0.03 & 0 \\ 0 & 0.43 & 0.19 & 0 & 0 & 0.15 & 0.03 & 0 \\ 0.43 & 0 & 0 & 0.19 & 0.15 & 0 & 0 & 0.03 \\ 0.07 & 0 & 0 & 0.03 & 0.03 & 0 & 0 & 0 \\ 0 & 0.07 & 0.03 & 0 & 0 & 0.03 & 0 & 0 \end{pmatrix}$$

which is a lower bound on, and a plausibly good approximation to, the probabilities of the 7-round differential (α_i, α_j). One sees that the differential (α_1, α_2) has probability about 1.22×2^{-58} and appears to be the largest 7-round differential probability. Our lower bound on the complexity of differential cryptanalysis shows then that at least 2^{59} encryptions will be required. The detailed cryptanalysis given in the Appendix shows that in fact the differential cryptanalysis attack will require all 2^{64} possible encryptions.

5. Improved PES

PES can be modified to improve its security without violating the design principles [2] used for PES. The resulting modified cipher will be called Improved PES and denoted as IPES. The only essential modification is that *a different (and simpler) permutation of subblocks is used at the end of each of the first 7 rounds.* The software implementation of IPES is in fact more efficient than that of PES.

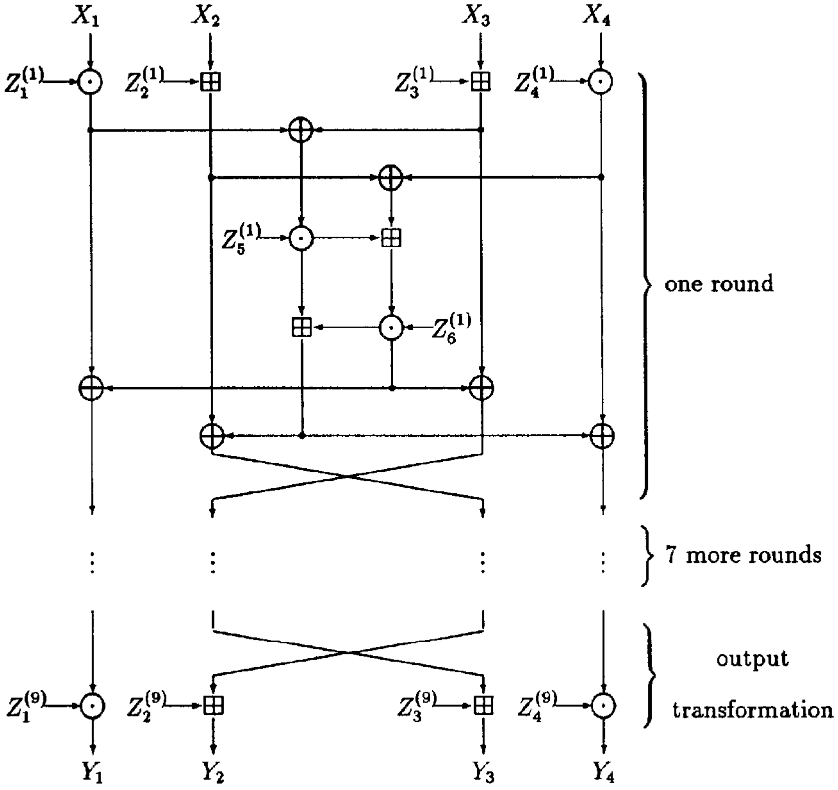
The computational graph of the encryption process of IPES is shown in Fig.4. Note that the permutation before the output transformation "undoes" the permutation at the end of 8-th round, i.e., at the end of 8-th round, the subblocks are *not* in fact permuted.

The key schedule used to generate the encryption key subblocks for IPES is the same as for PES (see [2]).

The decryption key DK for IPES is computed from the encryption key Z as follows,

$$\begin{aligned} &\text{for } r = 2,..,8: && (DK_1^{(r)}, DK_2^{(r)}, DK_3^{(r)}, DK_4^{(r)}) = (Z_1^{(10-r)^{-1}}, -Z_3^{(10-r)}, -Z_2^{(10-r)}, Z_4^{(10-r)^{-1}}) \\ &\text{for } r = 1,9: && (DK_1^{(r)}, DK_2^{(r)}, DK_3^{(r)}, DK_4^{(r)}) = (Z_1^{(10-r)^{-1}}, -Z_2^{(10-r)}, -Z_3^{(10-r)}, Z_4^{(10-r)^{-1}}) \\ &\text{for } r = 1,..,8: && (DK_5^{(r)}, DK_6^{(r)}) = (Z_5^{(9-r)}, Z_6^{(9-r)}), \end{aligned}$$

where Z^{-1} denotes the multiplicative inverse (modulo $2^{16}+1$) of Z, i.e., $Z \odot Z^{-1} = 1$ and where $-Z$ denotes the additive inverse (modulo 2^{16}) of Z, i.e., $-Z \boxplus Z = 0$. Thus, symmetry of encryption and decryption, which was one of the design principles of PES, is maintained in the sense that $IPES(IPES(X, Z), DK) = X$.



X_i : 16-bit plaintext subblock
Y_i : 16-bit ciphertext subblock
$Z_i^{(r)}$: 16-bit key subblock
$\oplus$: bit-by-bit exclusive-OR of 16-bit subblocks
$\boxplus$: addition modulo 2^{16} of 16-bit integers
$\odot$: multiplication modulo $2^{16} + 1$ of 16-bit integers with the zero subblock corresponding to 2^{16}

Figure 4: Encryption process of IPES.

Because of the involution property of the round function for PES, the transition probability matrix Π of round differentials of PES is symmetric. Thus, the one-round differentials (A,B) and (B,A) will have the same probability $P(B|A)$. For a highly likely one-round differential (A,B), the probability of a 2i-round differential (A,A) can thus be approximated by the probability of the corresponding 2i-round characteristics $(A, B, A, B, ..., A, B, A)$, i.e., $P_{2i}(A|A) \approx [P(B|A)]^{2i}$. Similarly, the probability of the 2i+1 round differential (A,B) can be approximated by $P_{2i+1}(B|A) \approx [P(B|A)]^{2i+1}$. For example, the probability of the plausibly most probable one-round differential for PES is $P(\Delta Y(1) = (0,0,0,1)|\Delta X = (0,0,0,-1)) \approx 2^{-9}$, [$-1$ stands for the integer $2^{16} - 1$,] the corresponding 7-round characteristics has probability about 2^{-61}. This suggests that $(\Delta X = (0,0,0,-1), \Delta Y(7) = (0,0,0,1))$ is the most likely 7-round differential. The analysis given in Appendix shows that it is indeed the (plausibly) most likely 7-round differential with probability about 2^{-58}, which is quite close to the probability of the corresponding characteristic.

The previous discussion suggests the following *design principle* for a Markov cipher: *The transition probability matrix of a Markov cipher should* not *be symmetric.*

Otherwise, the concatenation of the most probable one-round differential with itself r-1 times will tend to provide an (r-1)-round differential with high probability. The change of the permutation of subblocks between rounds of PES that is used in IPES is in accordance with this new design principle, i.e., the transition probability matrix of IPES is not symmetric. The change also significantly reduces the probabilities of the highly likely one-round differentials. For IPES, the (plausibly) most probable 1-round differential has probability $P(\Delta Y(1) = (1,0,1,0)|\Delta X = (1,1,0,0)) \approx 2^{-18}$,
and the (provably) most probable one-round differentials with $\Delta X = (1,0,1,0)$ and $\Delta X = (1,1,-1,1)$ are

$P(\Delta Y(1) = (1,1,-1,1)|\Delta X = (1,0,1,0)) \approx 2^{-34}$, and
$P(\Delta Y(1) = (1,1,0,0)|\Delta X = (1,1,-1,1)) \approx 2^{-34}$.

These figures imply that the corresponding 3-round characteristic $((1,1,0,0), (1,0,1,0), (1,1,-1,1), (1,1,0,0))$ has probability about 2^{-86}. This suggests that the probability of the 3-round differential $(\Delta X = (1,1,0,0), \Delta Y(3) = (1,1,0,0))$ should not be significantly larger than the average value, 2^{-64}. Other known one-round differentials for IPES with high probabilities are

$P(\Delta Y(1) = (1,H,0,3)|\Delta X = (1,H,H,1)) \approx 2^{-30}$,
$P(\Delta Y(1) = (1,0,H,3)|\Delta X = (1,H,0,3)) \approx 2^{-30}$,
$P(\Delta Y(1) = (1,H,H,1)|\Delta X = (1,0,H,3)) \approx 2^{-30}$,

where H stands for the integer 2^{15} whose corresponding 16-bit subblock is $(1,\overbrace{0,..,0,0}^{15bits})$. To date, we have found no evidence that there are any 3-round differentials for IPES whose probabilities are significantly larger than $2^{-m} = 2^{-64}$.

Finally, we remark that if one uses the difference defined as $DX = X_1 \bigoplus X_2$, then the most probable differentials (that we have found) for IPES become

$P(DY(1) = (0, H, 0, H) | DX = (0, H, H, 0)) = P(Z_5^{(1)} \in \{0, 1\}) = 2^{-15}$,
$P(DY(1) = (0, 0, H, H) | DX = (0, H, 0, H)) = P(Z_4^{(1)} \in \{0, 1\}) = 2^{-15}$, and
$P(DY(1) = (0, H, H, 0) | DX = (0, 0, H, H)) = P(Z_4^{(1)} \in \{0, 1\}, Z_5^{(1)} \in \{0, 1\}) = 2^{-30}$.
However, IPES is not a Markov cipher for this notion of difference. If DX is used as the definition of difference, the hypothesis of stochastic equivalence does not hold at all for IPES so that the differential probabilities computed for this notion of difference have no relation to an attack by differential cryptanalysis. The fact that the 3-round differential $(DX = (0, H, H, 0), DY(3) = (0, H, H, 0))$ for the difference DX has probability much larger than 3-round differentials for the "appropriate" difference ΔX used above has thus no significance for differential cryptanalysis.

Appendix: Detailed Differential Cryptanalysis of PES

1. Some One Round Differentials for PES

We first calculate the probabilities of certain one-round differentials for PES for pairs of input blocks that differ by a given value. This will enable us to calculate the probability of a $7\frac{1}{2}$-round differential from which it is usually possible to find the sub-key used in the last round.

Clearly a 16-bit number is its own inverse under the group operation $\oplus$. Let $-z$ denote the inverse of z under $\boxplus$, and z^{-1} the inverse of z under $\odot$. For any n-bit number z, let z' denote the n-bit complement of z. We also introduce some notation for the difference of two 16-bit numbers z_1, z_2 under the group operations $\odot$ and $\boxplus$. Let δ denote the difference under $\odot$ and ∂ denote the difference under $\boxplus$, i.e.,

$$\delta z = z_1 \odot z_2^{-1}, \qquad \partial z = z_1 \boxplus - z_2.$$

Then, for any 16-bit number k,

$$(z_1 \odot k) \odot (z_2 \odot k)^{-1} = \delta z, \qquad (z_1 \boxplus k) \boxplus - (z_2 \boxplus k) = \partial z.$$

Suppose $\delta z = z_1 \odot z_2^{-1} = 0$, where we recall that 2^{16} is represented by the integer 0 for the operation $\odot$. Then for $z_1, z_2 \notin \{0, 1\}$,

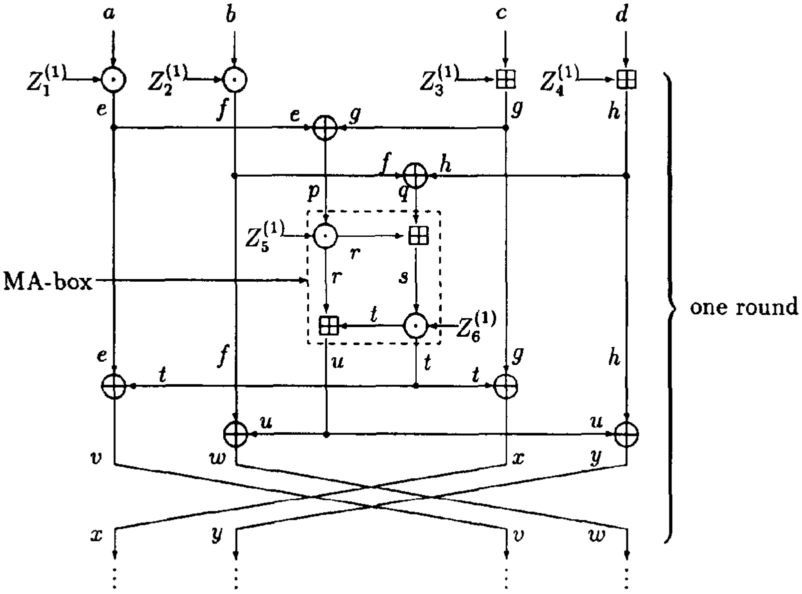
$$z_1 = 2^{16} z_2 = (2^{16} + 1) z_2 - z_2 = -z_2 \pmod{2^{16} + 1}, \text{ so } \quad z_1 + z_2 = 0 \pmod{2^{16} + 1.)}$$

Because z_1 and z_2 are positive 16-bit numbers, we have $z_1 + z_2 = 2^{16} + 1$. If $z_1 = 0$, then $z_2 = 1$, so in either case $z_1 + z_2 = 1 \pmod{2^{16}}$. Clearly, the converse also holds and so we have

$$\delta z = 0 \iff z_1 + z_2 = 1 \pmod{2^{16}}. \qquad (A.1)$$

Consider the MA-box, as defined in [2] and shown within the dashed lines in Figure 5. Suppose we have inputs (p_1, q_1) and (p_2, q_2) with outputs (t_1, u_1) and (t_2, u_2) respectively. The notion of the other 16-bit subblocks within the MA-box are defined in Figure 5. Suppose further that

$$p_1 + p_2 = 1 \pmod{2^{16}}, \qquad q_1 + q_2 = 0 \pmod{2^{16}}.$$



⊕ : bit-by-bit exclusive-OR of 16-bit subblocks
⊞ : addition modulo 2^{16} of 16-bit integers
⊙ : multiplication modulo $2^{16}+1$ of 16-bit integers
with the zero subblock corresponding to 2^{16}

Figure 5: The first round of PES and the notation used for differential cryptanalysis

Then $\delta p = 0$, so $\delta r = 0$ and hence $r_1 + r_2 = 1 \pmod{2^{16}}$. Thus,

$$s_1 + s_2 = (r_1 \boxplus q_1) + (r_2 \boxplus q_2) = r_1 + r_2 + q_1 + q_2 = 1 \pmod{2^{16}},$$

and hence $\delta s = 0$. Therefore $\delta t = 0$ so that $t_1 + t_2 = 1 \pmod{2^{16}}$ and hence

$$u_1 + u_2 = (r_1 \boxplus t_1) + (r_2 \boxplus t_2) = r_1 + r_2 + t_1 + t_2 = 2 \pmod{2^{16}}.$$

Thus, we have shown the following relationship between a pair of inputs and a pair of outputs of the MA-box,

$$p_1 + p_2 = 1, \;\; q_1 + q_2 = 0 \pmod{2^{16}} \quad \Rightarrow \quad t_1 + t_2 = 1, \;\; u_1 + u_2 = 2 \pmod{2^{16}}. \quad (A.2)$$

Consider one round of the cipher, as shown in Figure 5, where (a, b, c, d) are input subblocks and (x, y, v, w) the output subblocks. The intermediate results are defined in Figure 5. Suppose we have a pair of inputs with

$$(\delta a, \delta b, \partial c, \partial d) = (0, 0, 0, n) \quad \text{where } n \in S = \{\pm 1, \pm 3, \pm 5, \pm 7\}.$$

Then trivially we have $(\delta e, \delta g, \partial f, \partial h) = (0,0,0,n)$. Apart from the MA-box, an encryption round separates into two parts, and we consider first the half starting with (a,c). Suppose that

$$e_1 = (\alpha, 10...0, \theta)$$

for some $[16-(l+1)]$-bit number α, where $l \in \{0,\cdots,15\}$ and $\theta \in \{0,1\}$, so there are $(l-1)$ consecutive zeros before θ. Such an e_1 has probability 2^{-l}. From $(A.1)$, we know that, since $\delta e = 0$, $\quad e_1 + e_2 = 1 \pmod{2^{16}}$, and thus

$$e_2 = (\alpha', 10...0, \theta').$$

Now we can write g_i as

$$g_1 = g_2 = (\beta, 00...0, \phi),$$

where β is a $[16-(l+1)]$-bit number and $\phi \in \{0,1\}$. For a given l such an e_2 occurs with probability 2^{-l}. Then we have

$$\begin{aligned} p_1 &= e_1 \oplus g_1 = (\alpha \oplus \beta, 10...0, \theta \oplus \phi) \\ p_2 &= e_2 \oplus g_2 = (\alpha' \oplus \beta, 10...0, \theta' \oplus \phi) \end{aligned}$$

and thus $p_1 + p_2 = 1 \pmod{2^{16}}$. From the MA-box result $(A.2)$, it follows that

$$q_1 + q_2 = 0 \pmod{2^{16}} \quad \Rightarrow \quad t_1 + t_2 = 1 \pmod{2^{16}}.$$

Thus, with probability 2^{-l},

$$t_1 = (\gamma, 10...0, \rho), \qquad t_2 = (\gamma', 10...0, \rho'),$$

for some $[16-(l+1)]$-bit number γ and $\rho \in \{0,1\}$. Thus, for a given l, e_i, g_i and t_i all have the forms specified above with probability 2^{-3l}. Hence, e_i, g_i and t_i all have the forms specified above for some l with probability

$$\sum_{l=1}^{15} 2^{-3l} \approx \frac{1}{7}.$$

Thus, with probability $\frac{1}{7}$, we simultaneously obtain two results. Firstly, we have

$$\begin{aligned} v_1 &= e_1 \oplus t_1 = (\alpha \oplus \gamma, 00...0, \theta \oplus \rho) \\ v_2 &= e_2 \oplus t_2 = (\alpha \oplus \gamma, 00...0, \theta \oplus \rho) \end{aligned}$$

so that $v_1 = v_2$ and hence $\partial v = 0$. Secondly,

$$\begin{aligned} x_1 &= g_1 \oplus t_1 = (\beta \oplus \gamma, 10...0, \phi \oplus \rho) \\ x_2 &= g_2 \oplus t_2 = (\beta \oplus \gamma', 10...0, \phi \oplus \rho') \end{aligned}$$

so that $\quad x_1 + x_2 = 1 \pmod{2^{16}}$ and hence $\delta x = 0$. Thus, we have shown that if $q_1 + q_2 = 0 \pmod{2^{16}}$, then

$$(\delta a, \partial c) = (0,0) \Rightarrow (\delta x, \partial v) = (0,0) \quad \text{with probability} \approx \frac{1}{7}.$$

For the other half of the encryption process, we have $\delta f = 0$ and $\partial h \in S$. If we can find f and h such that $q_1 + q_2 = 0 \pmod{2^{16}}$, then, provided $\delta p = 0$, we have $u_1 + u_2 = 2 \pmod{2^{16}}$. In order to find a differential with high probability, we need to find such (u_1, u_2) with $\delta y = 0$ and $\partial w \in S$, where

$$y_i = h_i \oplus u_i, \qquad w_i = f_i \oplus u_i.$$

Thus, we need to find f, h and u, where

$$f_1 + f_2 = 1 \pmod{2^{16}}, \quad h_1 - h_2 \in S, \quad u_1 + u_2 = 2 \pmod{2^{16}}, \tag{A.3}$$

that satisfy, for q, w and y defined as above,

$$q_1 + q_2 = 0 \pmod{2^{16}}, \quad w_1 - w_2 \in S, \quad y_1 + y_2 = 1 \pmod{2^{16}}. \tag{A.4}$$

We can find most of the possible solutions fairly easily. Suppose

$$\begin{aligned} f_1 &= (\alpha, \theta_1), & f_2 &= (\alpha', \theta_2), \\ h_1 &= (\beta, \phi_1), & h_2 &= (\beta, \phi_2), \\ u_1 &= (\gamma, \rho_1), & u_2 &= (\gamma', \rho_2), \end{aligned}$$

for 12-bit numbers α, β, γ, and 4-bit numbers θ_i, ϕ_i, ρ_i. Then, from (A.3), we have

$$\theta_1 + \theta_2 = 17, \quad \phi_1 - \phi_2 \in S, \quad \rho_1 + \rho_2 = 18. \tag{A.5}$$

Now, by definition,

$$\begin{aligned} q_1 &= (\alpha \oplus \beta, \theta_1 \oplus \phi_1), & q_2 &= (\alpha' \oplus \beta, \theta_2 \oplus \phi_2), \\ y_1 &= (\beta \oplus \gamma, \phi_1 \oplus \rho_1), & y_2 &= (\beta \oplus \gamma', \phi_2 \oplus \rho_2), \\ w_1 &= (\alpha \oplus \gamma, \theta_1 \oplus \rho_1), & w_2 &= (\alpha \oplus \gamma, \theta_2 \oplus \rho_2), \end{aligned}$$

and thus (A.4) requires us to find solutions to

$$\begin{aligned} (\theta_1 \oplus \phi_1) + (\theta_2 \oplus \phi_2) &= 16 \\ (\phi_1 \oplus \rho_1) + (\phi_2 \oplus \rho_2) &= 17 \\ (\theta_1 \oplus \rho_1) - (\theta_2 \oplus \rho_2) &\in S. \end{aligned} \tag{A.6}$$

There are 514 triplets of 4-bit numbers satisfying (A.5) and (A.6). The numbers of solutions that correspond to each pair of elements of S are given in the matrix M below, where the rows and columns are in the order $\{-1, 1, -3, 3, -5, 5, -7, 7\}$,

$$M = \begin{pmatrix} 0 & 73 & 36 & 0 & 0 & 27 & 2 & 0 \\ 73 & 0 & 0 & 36 & 27 & 0 & 0 & 2 \\ 36 & 0 & 0 & 0 & 18 & 0 & 0 & 9 \\ 0 & 36 & 0 & 0 & 0 & 18 & 9 & 0 \\ 0 & 27 & 18 & 0 & 0 & 0 & 0 & 0 \\ 27 & 0 & 0 & 18 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 9 & 0 & 0 & 0 & 0 \\ 0 & 2 & 9 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This gives a matrix of transition probabilities, $T_{b,d} = 2^{-12}M$, where

$$T_{b,d} = 10^{-7}\begin{pmatrix} 0 & 178223 & 87891 & 0 & 0 & 65918 & 4882 & 0 \\ 178223 & 0 & 0 & 87891 & 65918 & 0 & 0 & 4882 \\ 87891 & 0 & 0 & 0 & 43945 & 0 & 0 & 21973 \\ 0 & 87891 & 0 & 0 & 0 & 43945 & 21973 & 0 \\ 0 & 65918 & 43945 & 0 & 0 & 0 & 0 & 0 \\ 65918 & 0 & 0 & 43945 & 0 & 0 & 0 & 0 \\ 4882 & 0 & 0 & 21973 & 0 & 0 & 0 & 0 \\ 0 & 4882 & 21973 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can now make a similar statement to the one above for the other half of the cipher. Given $p_1 + p_2 = 1 \pmod{2^{16}}$, $(\delta b, \partial d) = (0, n_1) \Rightarrow (\delta y, \partial w) = (0, n_2)$ for $n_1, n_2 \in S$ with probability given by the appropriate entry of $T_{b,d}$.

We can now calculate some approximate 1-round differential probabilities for PES. These are given by the transition matrix $T_{a,b,c,d} = \frac{1}{7}T_{b,d}$. Therefore,

$$T_{a,b,c,d} = 10^{-7}\begin{pmatrix} 0 & 25460 & 12556 & 0 & 0 & 9417 & 698 & 0 \\ 25460 & 0 & 0 & 12556 & 9417 & 0 & 0 & 698 \\ 12556 & 0 & 0 & 0 & 6278 & 0 & 0 & 3139 \\ 0 & 12556 & 0 & 0 & 0 & 6278 & 3139 & 0 \\ 0 & 9417 & 6278 & 0 & 0 & 0 & 0 & 0 \\ 9417 & 0 & 0 & 6278 & 0 & 0 & 0 & 0 \\ 698 & 0 & 0 & 3139 & 0 & 0 & 0 & 0 \\ 0 & 698 & 3139 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and hence

$$(\delta a, \delta b, \partial c, \partial d) = (0,0,0,n_1) \Rightarrow (\delta x, \delta y, \partial v, \partial w) = (0,0,0,n_2) \quad \text{for } n_1, n_2 \in S$$

with probability given by the appropriate entry of $T_{a,b,c,d}$.

It is of interest to know how accurate our approximation is. For purposes of comparison, the differential probabilities were calculated by simulation. The results, based on 10,000,000 randomly chosen input pairs for each element of S, are given in the matrix $\hat{T}$ given below:

$$\hat{T} = 10^{-7}\begin{pmatrix} 0 & 25291 & 12891 & 0 & 0 & 9755 & 817 & 0 \\ 25691 & 0 & 0 & 12769 & 9509 & 0 & 0 & 807 \\ 12712 & 0 & 0 & 0 & 6353 & 0 & 0 & 3154 \\ 0 & 12800 & 0 & 0 & 0 & 6324 & 3050 & 0 \\ 0 & 9482 & 6422 & 0 & 0 & 0 & 0 & 0 \\ 9396 & 0 & 0 & 6329 & 0 & 0 & 0 & 0 \\ 770 & 0 & 0 & 3148 & 0 & 0 & 0 & 0 \\ 0 & 757 & 3173 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

3. A Possible Cryptanalysis of the PES

We can now calculate a 7-round transition matrix, whose entries give the probability that, given that the input pair to the first round differ by a given value, then the ouput differences of the seventh round and the input differences to all the intermediate rounds are all of the required form. We denote this matrix by $T7_{a,b,c,d}$. It is easily calculated since

$$T7_{a,b,c,d} \;=\; T^7_{a,b,c,d} \;=\; 2^{-12\times 7}7^{-7}M^7.$$

The calculation gives the following matrix:

$$T7_{a,b,c,d} = \begin{pmatrix} 1.22\times 2^{-58} & 1.06\times 2^{-59} & 0.86\times 2^{-59} & 1.12\times 2^{-62} \\ 1.06\times 2^{-59} & 0.92\times 2^{-60} & 1.52\times 2^{-61} & 0.96\times 2^{-63} \\ 0.86\times 2^{-59} & 1.52\times 2^{-61} & 1.21\times 2^{-61} & 0.96\times 2^{-63} \\ 1.12\times 2^{-62} & 0.96\times 2^{-63} & 0.96\times 2^{-63} & 1.76\times 2^{-65} \end{pmatrix}$$

where the (i,j) entry of the matrix gives a plausibly good approximation of the probability of the 7-round differentials of the form

$$(\delta a^1, \delta b^1, \partial c^1, \partial d^1) = (0,0,0,\pm(2i-1)), \quad (\delta a^7, \delta b^7, \partial c^7, \partial d^7) = (0,0,0,\mp(-1)^{i+j}(2j-1))$$

where the superscript indicates the round index.

We can now obtain the key as follows. Suppose initially we choose $(\delta a^1, \delta b^1, \partial c^1, \partial d^1) = (0,0,0,1)$, then we know that after 7 encryption rounds that

$$(\delta a^8, \delta b^8, \partial c^8, \partial d^8) = (\delta e^8, \delta f^8, \partial g^8, \partial h^8) = (0,0,0,-1)$$

with probability 1.22×2^{-58}. From our earlier arguments on the (a,c) half of the cipher, it follows that $(\delta a, \partial c) = (0,0) \Rightarrow \delta p = 0$ with probability $\approx \frac{1}{3}$, and, by an argument similar to one given above of searching all the 4-bit numbers, we can show that

$$(\delta b, \partial d) = (0,1) \Rightarrow q_1 + q_2 = 0 \pmod{2^{16}}$$

with probability $\approx \frac{42}{256}$. Hence we have

$$t_1 + t_2 = 1 \pmod{2^{16}}, \quad u_1 + u_2 = 2 \pmod{2^{16}},$$

with probability $\approx \frac{14}{256} = 1.75\times 2^{-5}$. After 7 rounds, we have

$$\begin{gathered}(\delta e^8, \delta f^8, \partial g^8, \partial h^8) = (0,0,0,-1) \\ p_1^8 + p_2^8 = 1, \;\; q_1^8 + q_2^8 = 0, \;\; t_1^8 + t_2^8 = 1, \;\; u_1^8 + u_2^8 = 2\end{gathered}$$

with probability 1.07×2^{-62}.

We can now determine the 96 key bits that are used from this stage to the ciphertext in the following way. If we denote the ciphertext as $(e_i^9, f_i^9, g_i^9, h_i^9)$, then observing one ciphertext pair gives us

$$(\delta e^9, \delta f^9, \partial g^9, \partial h^9) \;=\; (\delta x^8, \delta y^8, \partial v^8, \partial w^8).$$

As before, we divide the encryption round into two halves and we consider first the half ending with (x, v). For a given $(\delta x^8, \partial v^8)$, each 48-bit triplet (t_1^8, x_1^8, v_1^8) determines all the other quantities in that half of the cipher outside the MA-box. We have additionally to satisfy three 16-bit constraints in order to satisfy the conditions for the $7\frac{1}{2}$-round characteristics given above. Hence, for a given $(\delta x^8, \partial v^8)$, there will on the average be one 48-bit triplet (t_1^8, x_1^8, v_1^8) and hence also one possible value for the key blocks $Z_1^{(9)}$ and $Z_3^{(9)}$ (in the notation of Figure 2). However, some values of $(\delta x^8, \partial v^8)$ will give considerably more triplets (t_1^8, x_1^8, v_1^8). For example, if $(\delta x^8, \partial v^8) = (0, 0)$, the differential ouput, then one seventh of all triplets (t_1^8, x_1^8, v_1^8) will be possible. Such differences do occur infrequently, (the one given above with probability 2^{-32}) and do not significantly affect the argument given below. Note that these 48-bit triplets and two 16-bit key blocks can be pre-calculated for each value of $(\delta e^9, \partial g^9)$. A similar result obviously holds for $(\delta f^9, \partial h^9)$. Combining the two results, we can see that, for each value of $(\delta e^9, \delta f^9, \partial g^9, \partial h^9)$, we obtain on the average one value for $(p_1^8, q_1^8, t_1^8, u_1^8)$, and from inverting the MA-box, we then obtain on average one value for the key blocks $Z_5^{(8)}$ and $Z_6^{(8)}$. Hence, for each plaintext-ciphertext pair, we obtain on the average one possible value for 96 key bits, that is to say, a particular value for the 96 key bits occurs with a probability of 2^{-96} per encryption pair.

If a key occurs with probability p, then in 2^N encryption pairs, the key occurs k times with probability

$$\binom{2^N}{k} p^k (1-p)^{2^N - k} \approx \frac{(2^N p)^k}{k!} e^{-2^N p}.$$

Thus, in 2^N encryption pairs, an incorrect key will occur two or more times with probability $\frac{1}{2} \times 2^{2(N-96)}$. If we encrypt the whole message space ($N = 64, 2^N p = 2^{-32}$), then a wrong key will occur two or more times with probability

$$1 - P(\text{wrong key occurs 0 or 1 time}) = 1 - exp(-2^{-32}) - 2^{-32} exp(-2^{-32}) \approx \frac{1}{2} 2^{-64},$$

so this event will happen for 2^{31} of the 96-bit keys. The correct key, however, occurs with this probability whenever the $7\frac{1}{2}$-round differential does not occur, and with probability $p = 1.07 \times 2^{-62}$ when the differential does occur. Thus in 2^{64} encryption pairs ($2^N p = 4.28$), the correct key will occur less than twice with probability

$$exp(-4.28) + 4.28 exp(-4.28) \approx 0.073,$$

so the correct key is highly likely to occur more than once (about 93% of the time). To find the key, we can then just try all the 96-bit keys that occur more than once in this procedure, there are approximately 2^{31} such keys. However, since the subkeys are determined by 128-bit key, there are 2^{32} keys that give rise to each last round 96-bit subkey, so there are 2^{63} keys that give last round subkeys occurring more than once. We therefore have a reduced key search of about 2^{63} after all the encryptions have taken place. Since the true key occurs with a much larger probability than any of the false subkeys, we would expect to deal with it before we had tried too many false keys. In that way, we should have to try only the 2^{32} keys that give the correct 96-bit subkey, and a few other subkeys perhaps. Thus, the key search will in practice be reduced to almost 2^{32}.

4. Conclusions

The cryptanalysis of PES given above is, of course, computationally infeasible, but it does illustrate some interesting points. The first is that the true strength of the standard PES algorithm is of the order 2^{64} encryptions, a considerable reduction from the work that a cryptanalyst would expected in an exhaustive key search for the 128-bit key. Second, it shows that a chosen plaintext attack would be computationally possible on a reduced round standard PES. It can be seen that the attack outlined above works on an m-round PES with roughly 2^{8m} encryption pairs. Thus a 4-round PES could be broken with roughly 2^{32} encryption pairs. These conclusions of course do not apply to the modified PES, called IPES and described in this paper; IPES appears to be invulnerable to differential cryptanalysis for the reason given in Section 5.

Acknowledgements

The authors are grateful to Prof. Hans Bühlmann of the ETH Zürich for a helpful suggestion. This research was supported by the Swiss Commission for the Advancement of Scientific Research, Research Grant KWF 2146.1, and by the British Science and Engineering Research Council, Research Grant GR/E 64640.

References

[1] E.Biham and A.Shamir, "Differential Cryptanalysis of DES-like Cryptosystems", to appear in Journal of Cryptology, Vol.4, No.1, 1991.

[2] X.Lai and J.L.Massey, "A Proposal for a New Block Encryption Standard", Advances in Cryptology-EUROCRYPT'90, Springer-Verlag, Berlin 1991, pp. 389-404.