Pseudorandomness and Maximum Average of Differential Probability of Block Ciphers with SPN-Structures like E2

Makoto Sugita[1] and Kazukuni Kobara[2] and Hideki Imai[2]
[1] NTT Wireless Systems Laboratories
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa-ken, 239-0847 Japan
E-mail: sugita@pcs.wslab.ntt.co.jp
[2] Institute of Industrial Sciences, The University of Tokyo
Roppongi, Minato-ku, Tokyo 106-8558, Japan
E-mail:{kobara, imai}@imailab.iis.u-tokyo.ac.jp

Abstract

This paper introduces a new estimation method of Luby-Rackoff's pseudorandomness and maximum average of differential probability of block ciphers with SPN(Substitution and Permutation Network)-structures like E2. In this paper, we analyze the pseudorandomness of the SPN-structure and E2-like transformations and show that this can be easily calculated by simple matrix calculation, and clarify that the linear transformation used in E2 offers good pseudorandomness. Moreover, we examine the maximum average of the differential probability of the SPN-structure. We show that this can be calculated recursively by a novel calculation method and confirm that the linear transformation used in E2 offers good immunity for differential attacks when used in the 4-round SPN-structure.

keywords. E2, E2-like transformation, SPN-structure, maximum average of differential probability, pseudorandomness

1 Introduction

In this paper, we analyze security of block ciphers with SPN(Substitution and Permutation Network)-structures like E2. We consider two definitions of security - Luby-Rackoff's pseudorandomness and maximum average of differential probability. We introduce a new estimation method of Luby-Rackoff's pseudorandomness and maximum average of differential probability of block ciphers with SPN(Substitution and Permutation Network)-structures like E2.

The notion of a pseudorandom function generator (PRFG) was introduced by Goldreich, Goldwasser and Micali in [GGM84] who showed how to efficiently construct a pseudorandom function generator from a pseudorandom bit generator. In [LR86], Luby and Rackoff defined a pseudorandom invertible permutation generator (PRPG). Using ideas behind the design of the Data Encryption Standard, they showed how to efficiently construct a pseudorandom invertible permutation generator from an pseudorandom function generator. A practical implication of their result is that any pseudorandom function generator can be used to construct a block private key cryptosystem that is secure against chosen plaintext attack, which is one of the strongest known attacks against a cryptosystem. They also defined a generalized pseudorandom function, i.e. (n, m, k, ϵ) - pseudorandom function (PRF). They showed (n, m, k, ϵ)-PRF constructs $(2n, m, k, \epsilon')$-PRP for some ϵ', which implies that (n, m, k, ϵ)-PRP can also construct $(2n, m, k, \epsilon')$-PRP for some ϵ' by regarding (n, m, k, ϵ)-PRP as (n, m, k, ϵ'')-PRF for some ϵ''. These results imply that pseudorandomness can be used as a important measure of immunity against chosen plaintext attack even if the encrypting functions (s-boxes) constructing block cipher are bijective.

In [S97], we showed one sufficient condition such that the basic transformations with recursive structures yield PRF, and proved that the $(5, 3)$ and the $(5, 3, \cdots, 3)$-round iterations of the basic transformations of MISTY (proposed by Matsui in [Ma97]) satisfy this condition. They yield a PRF, while $(4, 3)$ and the $(4, 3, \cdots, 3)$-round iterations do not. In [S97-2], we showed stronger sufficient condition for the basic transformations to be PRF, and show that both the $(5, 3, \cdots, 3)$-round iteration of the basic transformations of MISTY and the $(4, 3, \cdots, 3)$-round iteration of the basic transformations of MISTY1 satisfy this condition, and as a result, yield PRF.

The block cipher E2 was proposed in [K98] as an AES candidate. This cipher uses Feistel structures as a global structure like DES, and uses the SPN(Substitution and Permutation Network)-structure in s-boxes. In this paper, we apply our previous condition to SPN-structures and basic transformations of E2, and show that this can be easily calculated by some matrix calculation, and clarify that the linear transformation used in E2 offers good pseudorandomness.

As another measure of the security for block ciphers, the maximum average of differential probability was defined by Nyberg and Knudsen by generalizing provable security against linear and differential cryptanalysis by Biham and Shamir [NK 94]. In this paper, we estimate the maximum average of the differential probability of the SPN-structure. In [K98], they state that this evaluation is practically impossible, but we show that this can be calculated recursively by a novel but simple calculation and showed that the linear transformation used in E2 has good property as it is used in the SPN-structure.

This paper is organized as follows.

We describe the pseudorandomness of block ciphers in Section 2.

In section 3, we describe SPN-structures and block cipher E2.

In section 4, Applying our sufficient condition, we analyze the pseudorandomness of the SPN-structure and E2-like transformations and show that pseudorandomness can be easily evaluated by the matrix calculation proposed herein, and clarify that the linear transformation used in E2 has good pseudorandomness as it is used in the 4-round SPN-structures.

In section 5, we estimate the maximum average of the differential probability of the SPN-structure.

2 Preliminary

2.1 Notation

For $s_1, s_2 \in \{0,1\}^n$, $s_1 \oplus s_2$ denotes the bit-wise XOR of s_1 and s_2. F^n denotes the set of all functions from $\{0,1\}^n$ to $\{0,1\}^n$. $F^n_{\mathcal{Z}}$ denotes the set of all functions from $\{0,1\}^n$ to $\{0,1\}^n$ with the key space $\mathcal{Z}$.

2.2 Pseudorandom Functions Generator

In this subsection, pseudorandom function generator (PRFG) is defined. We denote a random function $r : \{0,1\}^n \to \{0,1\}^n$ as a function that assigns to all arguments $x \in \{0,1\}^n$ independent and completely random values $r(x) \in \{0,1\}^n$. First we introduce a generalized random function for the proof of pseudorandomness of the basic transformation constructing block ciphers.

Definition 1 *A keyed function $r_z : \{0,1\}^n \to \{0,1\}^n (z \in \mathcal{Z})$ with the key space $\mathcal{Z}$ is a generalized random function if for every $x_1, x_2 \in \{0,1\}^n (x_1 \neq x_2)$ and $z_1, z_2 \in \mathcal{Z}$, $r_{z_1}(x_1)$ and $r_{z_2}(x_2)$ are random and jointly statistically independent. As a special case, conventional random functions of F^n are generalized random functions if regarded as functions with key space $\mathcal{Z}$ (In this case, the output value is not depend on the key value $z \in \mathcal{Z}$.).*

Next we introduce the condition $\delta(n)$-random for a random variable in order to prove the pseudorandomness of the basic transformations of the block ciphers.

Definition 2 *Let X be a random variable that takes on values $x \in \{0,1\}^n$, and (X, X) be a 2-dimensional random variable that takes on values $(x_1, x_2) \in (\{0,1\}^n)^2$.*

We define X as $\delta(n)$-random if for some event Δ, such that $P(\Delta) \leq \delta(n)$, (X, X) takes values randomly over the complementary event $\bar{\Delta} = (\{0,1\}^n)^2 - \Delta$.

[LR86] defined the PRFG. In the following three definitions, we omit the restriction on the function (which [LR86] denotes as distinguishing circuits) because, in [M92], they showed that it is not essential in the proof and can be omitted.

Definition 3 (LR86) *A family $\mathcal{F}_{\mathcal{Z}} = \{f_z : z \in \mathcal{Z}\}$ of functions $f_z : \{0,1\}^n \to \{0,1\}^m$ is an (n, m, k) pseudorandom function (PRF) with the key space $\mathcal{Z}$ if for every subset $\{x_1, ..., x_k\}$ of $\{0,1\}^n$, $f_z(x_1), ..., f_z(x_k)$ are uniformly distributed over $\{0,1\}^m$ and are jointly statistically independent, when z is randomly chosen from $\mathcal{Z}$.*

Definition 4 (LR86) *A family $\mathcal{F}_{\mathcal{Z}} = \{f_z : z \in \mathcal{Z}\}$ of functions $f_z : \{0,1\}^n \to \{0,1\}^m$ is an (n, m, k, ϵ) pseudorandom function (PRF) with key space $\mathcal{Z}$ if for all functions $g : (\{0,1\}^m)^k \to \{0,1\}$ and for every subset $\{x_1, ..., x_k\}$ of $\{0,1\}^n$, for z randomly chosen from $\mathcal{Z}$,*

$$|P[g(f_z(x_1), ..., f_z(x_k)) = 1] - P[g(r_1, ..., r_k) = 1]| \leq \epsilon$$

where $r_1, ..., r_k$ are independent and randomly chosen from $\{0,1\}^m$

Definition 5 (LR86) *A pseudorandom function generator (PRFG) with the key length function $l(n)$ and degree of local randomization $k(n)$ is the family*

$$\mathcal{F} = \{\mathcal{F}^n_{\{0,1\}^{l(n)}} : n \in \mathcal{N}\},$$

where $\mathcal{F}^n_{\{0,1\}^{l(n)}}$ is an $(n, n, k(n), \epsilon(n))$ PRF with key space $\{0,1\}^{l(n)}$ that is, for every given argument and key computable in time polynomial in n, independent of the number of previous evaluations, where $\epsilon(n)$ vanishes faster than $1/Q(n)$ for every polynomial $Q(n)$

[LR86] defined a pseudorandom invertible permutation generator as a family of permutations that is also a PRFG family, where the required security property is to approximate, as closely as possible, a random function. However, in [BKR98], they use another model for PRP of [Sh49], where the required security property is to approximate, as closely as possible, a random permutation. They also state that the two models of security for PRP are nearly the same when the number of encrypted blocks m is small, and that PRF is a better tool than PRP, from two points of view: it permits easier and more effective analysis of the designed scheme, and the resulting schemes have a greater level of proven quantative security. This leads us to suggest that for the purpose of protocol design, what we really want are PRFs, not PRPs. Therefore, in the following three definitions for PRPs, we use the models of [LR86].

Definition 6 (LR86) *A family $\mathcal{F}_{\mathcal{Z}} = \{f_z : z \in \mathcal{Z}\}$ of permutations $f_z : \{0,1\}^n \rightarrow \{0,1\}^n$ is an (n, k) pseudorandom permutation (PRP) with the key space $\mathcal{Z}$ if for every subset $\{x_1, ..., x_k\}$ of $\{0,1\}^n$, $f_z(x_1), ..., f_z(x_k)$ are uniformly distributed over $\{0,1\}^n$ and are jointly statistically independent, when z is randomly chosen from $\mathcal{Z}$.*

Definition 7 (LR86) *A family $\mathcal{F}_{\mathcal{Z}} = \{f_z : z \in \mathcal{Z}\}$ of permutations $f_z : \{0,1\}^n \rightarrow \{0,1\}^n$ is an (n, k, ϵ) pseudorandom permutation (PRP) with the key space $\mathcal{Z}$ if for all functions $g : (\{0,1\}^n)^k \rightarrow \{0,1\}$ and for every subset $\{x_1, ..., x_k\}$ of $\{0,1\}^n$, for z is randomly chosen from $\mathcal{Z}$,*

$$|P[g(f_z(x_1), ..., f_z(x_k)) = 1] - P[g(r_1, ..., r_k) = 1]| \leq \epsilon$$

where $r_1, ..., r_k$ are independent and randomly chosen from $\{0,1\}^n$

The existence of PRP under the assumption of the existence of PRF was proved in [LR86] and [M92].

Definition 8 (LR86) *A pseudorandom permutation generator (PRPG) with the key length function $l(n)$ and degree of local randomization $k(n)$ is the family*

$$\mathcal{F} = \{\mathcal{F}^n_{\{0,1\}^{l(n)}} : n \in \mathcal{N}\},$$

where $\mathcal{F}^n_{\{0,1\}^{l(n)}}$ *is an* $(n, k(n), \epsilon(n))$ *PRP with key space* $\{0,1\}^{l(n)}$ *that is for every given argument and key computable in time polynomial in* n*, independent of the number of previous evaluations, where* $\epsilon(n)$ *vanishes faster than* $1/Q(n)$ *for every polynomial* $Q(n)$

Note. The existence of PRPG under the assumption of the existence of PRFG is proved in [LR86] and [M92].

3 Block Cipher E2

3.1 SPN-Structures [K98]

In [K98], SPN-Structures are defined. First we define the 2-round SPN-structure as in Fig.1.

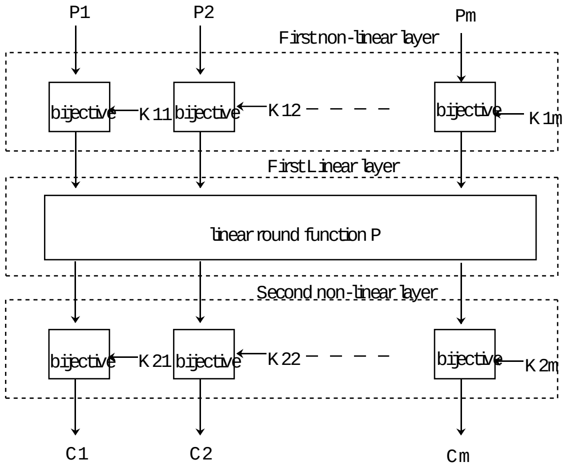

Figure 1: 2-round SPN-structure

This structure consists of two kinds of layers, i.e. non-linear layer and bijective linear layer. Each layer has the following feature.

Non-linear layer: This layer is composed of m parallel n-bit bijective s-boxes.

Linear layer: This layer is composed of bitwise XORs, where inputs are transformed linearly to outputs per byte (n-bits).

[K98] introduces a matrix expression $P_E = \{a_{ij}\}$ of linear round function E, where $a_{ij} = 1$ means that the input of i-th s-box in second nonlinear layer linearly depends on the output of j-th s-box in first nonlinear layer, and $a_{ij} = 0$ means does not.

Next we define the N-round SPN-structure as in Fig.2. This layer consists of $(2N-1)$ layers. First is the nonlinear layer, second linear layer, generally, i-th nonlinear layer $(i = 1, \cdots, N-1)$, and i-th linear layer $(i = 1, \cdots, N)$ in this order. Furthermore, for the functions $f_{11}, \cdots, f_{Nm}$, we denote N-round SPN-structures as $\mathrm{SPN}_{N,m}(f_{11}, f_{12}, \cdots, f_{Nm})$, where the functions f_{ij} correspond to the bijection s_{ij} in Fig.2 $(1 \le i \le N, 1 \le j \le m)$.

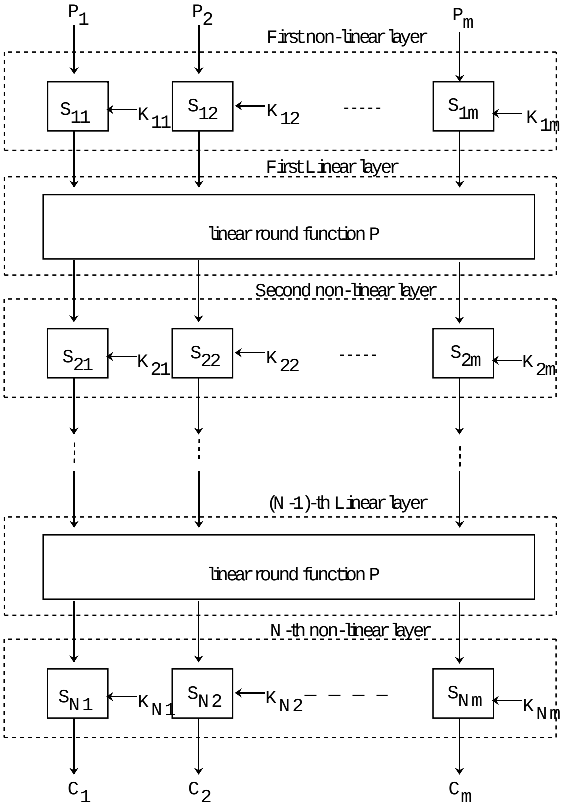

Figure 2: n-round SPN-structure

3.2 E2-like transformations

[K98] proposed the block cipher E2. This cipher has Feistel structures and its s-box is composed of the 2-round Feistel structures defined in the previous subsection. Here we define E2-like transformations as the Feistel structure with s-box composed of N-round (in this case, 2-round) SPN-structures as in Fig.3 . Furthermore, for the functions $f_{111}, \cdots, f_{sNm}$, we denote s-round E2 like transformations as $\mathrm{E2}_{s,N,m}(f_{111}, f_{112}, \cdots, f_{sNm})$, where the functions f_{ijk} correspond to the bijection s_{jk} in the i-th round s-box in Fig.2, Fig.3 $(1 \leq i \leq s, 1 \leq j \leq N, 1 \leq k \leq m)$.

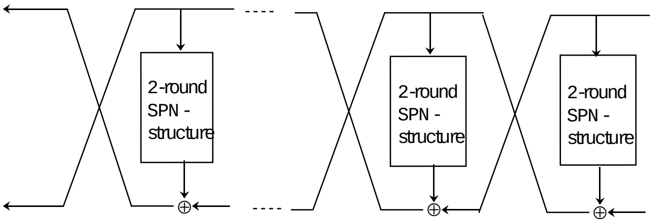

Figure 3: E2-like transformations

4 Pseudorandomness of SPN-structure and block cipher E2

4.1 Sufficient Condition for PRFG

In [S97] and [S97-2], we introduced effective sufficient conditions for the basic transformation of block ciphers to yield PRFG, and proved that basic transformations for MISTY(1) satisfy this condition for some round numbers, where MISTY is the block cipher proposed in [Ma97]. This condition for pseudorandomness can be applied to various types of block ciphers in AES candidates with SPN-structures or Feistel structures, where SPN-structures are used in CRYPTON, E2, LOKI97, MARS, RC6, RIJNDAEL, SAFER and SERPENT, and Feistel structures are used in CAST-256, DEAL, DFC, E2, LOKI97, MAGENTA.

Here we describe the condition in [S97-2].

Definition 9 *For a list of functions (oracle gates) $f_1, f_2, \cdots, f_s \in F^n$, let $f = \psi(f_1, f_2, \cdots, f_s) : (\{0,1\}^n)^m \to (\{0,1\}^n)^m$ be an acyclic circuit that consists of nbit-and/nbit-or/nbit-not/nbit-xor, nbit-fan-out, $f_i (i = 1, 2, \cdots, s)$, where f includes only one $f_i (i = 1, \cdots, s)$. f_i appears only once in f. Let $y_1 \bullet y_2 \bullet \cdots \bullet y_m \in (\{0,1\}^n)^m$ be an input of f, and let $z_1 \bullet z_2 \bullet \cdots \bullet z_m \in (\{0,1\}^n)^m$ be an output of f which is defined by $z_1 \bullet z_2 \bullet \cdots \bullet z_m = f(y_1 \bullet y_2 \bullet \cdots \bullet y_m)$. Let $IP_{f_a} \in \{0,1\}^n (a \in \{1, 2, \cdots, s\})$ be an input of f_a in the circuit f when the input of f is $y_1 \bullet y_2 \bullet \cdots \bullet y_m$, let $OP_{f_a} \in \{0,1\}^n$ be an output of f_a i.e. $OP_{f_a} = f_a(IP_{f_a})$ when the input of f is $y_1 \bullet y_2 \bullet \cdots \bullet y_m$.*

Let $y'_1 \bullet y'_2 \bullet \cdots \bullet y'_m \in (\{0,1\}^n)^m$ be another input of f, and let $z'_1 \bullet z'_2 \bullet \cdots \bullet z'_m \in (\{0,1\}^n)^m$ be an output of f which is defined by $z'_1 \bullet z'_2 \bullet \cdots \bullet z'_m = f(y'_1 \bullet y'_2 \bullet \cdots \bullet y'_m)$. Let $IP'_{f_a} \in \{0,1\}^n (a \in \{1, 2, \cdots, s\})$ be an input of f_a in the circuit f when the input of f is $y'_1 \bullet y'_2 \bullet \cdots \bullet y'_m$, let $OP'_{f_a} \in \{0,1\}^n$ be an output of f_a i.e. $OP'_{f_a} = f_a(IP'_{f_a})$ when the input of f is $y'_1 \bullet y'_2 \bullet \cdots \bullet y'_m$. Let $\mathcal{Z}$ be the key space.

We say ψ satisfies $\delta(n)$-condition 1′ if and only if there exist $(i_1, i_2, \cdots, i_m)$, $(j_1, j_2, \cdots, j_m)$, $i_a, j_b \in \{1, 2, \cdots, s\}$ $(a, b \in \{1, 2, \cdots, m\})$, that satisfy the following 4 conditions (cf. Fig.4).
1′.1) For every $a, b \in \{1, 2, \cdots, m\}$, $i_a \neq i_b (a \neq b)$, $j_a \neq j_b (a \neq b)$, $i_a \neq j_b$.
1′.2) When $f_1, f_2, \cdots, f_s \in F^n$, if $y_1 \neq y'_1$ then $IP_{f_{i_1}} \neq IP'_{f_{i_1}}$, and for every $a \in \{2, \cdots, m\}$, if $y_l = y'_l (l = 1, 2, \cdots, a-1)$ and $y_a \neq y'_a$ then $IP_{f_{i_a}} \neq IP'_{f_{i_a}}$.
1′.3) For every $a, b \in \{1, 2, \cdots, m\}$, if $OP_{f_{i_a}}$ is random and $f_1, \cdots, f_{i_a - 1}, f_{i_a + 1}, \cdots, f_s$ are random functions of F^n then $IP_{f_{j_b}}$ is $\delta(n)$-random, where we regard $OP_{f_{i_a}}$ and $IP_{f_{j_b}}$ as random variables. Also if $OP_{f_{i_a}}$ is random and $f_1, \cdots, f_{i_a - 1}$, $f_{i_a + 1}, \cdots, f_s$ are generalized random functions (oracle gates) of $F^n_{\mathcal{Z}}$, then $IP_{f_{j_b}}$ is $\delta(n)$-random.
1′.4) If $OP_{f_{j_1}}, OP_{f_{j_2}}, \cdots, OP_{f_{j_m}}$ are random and jointly statistically independent, then $z_1 \bullet z_2 \bullet \cdots \bullet z_m$ is random.

This definition is essentially composed of three relations: 1′.2) refers to the relations between the inputs $y_1, \cdots, y_m$ and the inputs of m input-related functions (oracle gates) $f_{i_1}, \cdots, f_{i_m}$. 1′.4) refers to the relations between

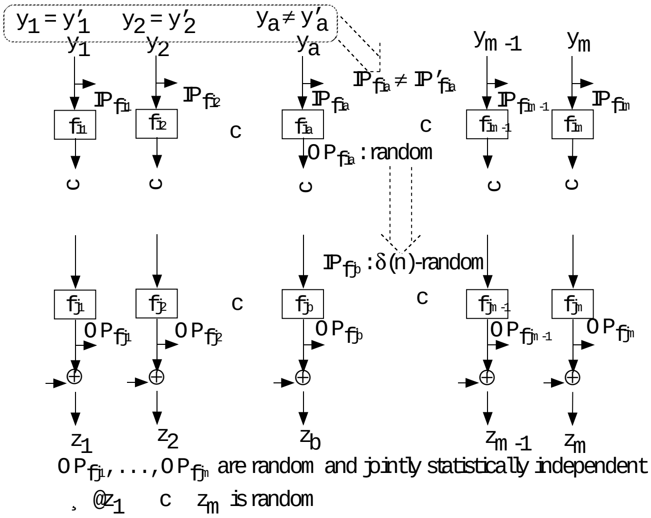

Figure 4: Definition of $\delta(n)$-condition 1'

the outputs $z_1, \cdots, z_m$ and the outputs of m output-related functions (oracle gates) $f_{j_1}, \cdots, f_{j_m}$. $1'.3$) refers to the relations between the outputs of m input-related functions (oracle gates) $f_{i_1}, \cdots, f_{i_m}$ and the inputs of m output-related functions $f_{j_1}, \cdots, f_{j_m}$. This definition is a generalization of the essence used in the proof of the pseudorandomness of DES-like transformation in [LR86].

The following lemma proves that the condition above implies PRFG.

Lemma 1 *For a list of functions $f_1, f_2, \cdots, f_s \in F^n$, let $f = \psi(f_1, f_2, \cdots, f_s)$: $(\{0,1\}^n)^m \to (\{0,1\}^n)^m$, be an acyclic circuit that consists of nbit-and/nbit-or/nbit-not/nbit-xor, nbit-fan-out, oracle gates f_i $(i = 1, 2, \cdots, s)$, where f includes only one $f_i (i = 1, \cdots, s)$. For every function g : $((\{0,1\}^n)^m)^k \to \{0,1\}$ and for every set of k arguments $x_1, \cdots, x_k$, if f satisfies $\delta(n)$-condition 1', then we have*

$$|P[g(f(x_1), \cdots, f(x_k)) = 1 : f \in_R \psi_n((F^n)^s)] - P_g| \leq mk^2(\delta(n)/2 + 2^{-n-1}).$$

Proof. Let $f_1, f_2, \cdots, f_s$ be random oracle gates (functions) of F^n, let $f = \psi(f_1, f_2, \cdots, f_s)$, let $x_l = y_{1l} \bullet y_{2l} \bullet \cdots \bullet y_{ml} \in (\{0,1\}^n)^m (1 \leq l \leq k)$, let $z_{1l} \bullet z_{2l} \bullet \cdots \bullet z_{ml} = f(x_l)(1 \leq l \leq k)$, let $(i_1, i_2, \cdots, i_m)$, $(j_1, j_2, \cdots, j_m)$ be the index used in conditions $1'.1$)-$1'.4$) in $\delta(n)$-condition $1'$, let $IP_{f_{j_a} l}$ be an input of f_{j_a} when the input of f is x_l, and let $OP_{f_{j_a} l}$ be the output of f_{j_a} when the input of f is x_l. We may, for the rest of the proof, assume without loss of generality that x_l, $1 \leq l \leq k$, are distinct because of the same reason as given in the Lemma 1 of [M 92].

For every $a \in \{1, 2, \cdots, m\}$, let $\mathcal{E}_{IP_{f_{j_a}}}$ denote the events that $IP_{f_{j_a} 1}, \cdots, IP_{f_{j_a} k}$ are distinct, and let $\mathcal{E}$ be the event that for every $a \in \{1, 2, \cdots, m\}$ $\mathcal{E}_{IP_{f_{j_a}}}$ occurs. If $\mathcal{E}_{IP_{f_{j_a}}}$ occurs, then $OP_{f_{j_a} 1}, OP_{f_{j_a} 2}, \cdots, OP_{f_{j_a} k}$ are random because f_{j_a} is a random function. Thus if $\mathcal{E}_{IP_{f_{j_a}}}$ occurs for all $a \in$

$\{1, 2, \cdots, m\}$, $f(x_1), f(x_2), \cdots, f(x_k)$ are random because of $1'.4)$ in $\delta(n)$-condition $1'$, and thus $f = \psi_n(f_1, f_2, \cdots, f_s)$ behaves precisely like a function chosen randomly from F^{mn} . Therefore the distinguishing probability is upper bounded by

$$|P[g(f(x_1), \cdots, f(x_k)) = 1 : f \in_R \psi((F^n)^m)] - P_g| \leq 1 - P[\mathcal{E}].$$

We now derive an upper bound on $1 - P[\mathcal{E}] = P[\bar{\mathcal{E}}]$, where $\bar{\mathcal{E}}$ denotes the complementary event of $\mathcal{E}$. $\bar{\mathcal{E}}$ is the union of the $m \left(\begin{array}{c} k \\ 2 \end{array} \right)$ events $\{IP_{f_{ja}}u = IP_{f_{ja}}v\}$ for $1 \leq u < v \leq k$, $1 \leq a \leq m$. The probability of the union of several events is upper bounded by the sum of the probabilities, and hence

$$1 - P[\mathcal{E}] = P[\bar{\mathcal{E}}] \leq \sum_{1 \leq a \leq m} \sum_{1 \leq u < v \leq k} P[IP_{f_{ja}}u = IP_{f_{ja}}v]. \qquad (1)$$

For $u \neq v$ we have

$$P[IP_{f_{ja}}u = IP_{f_{ja}}v] \leq \delta(n) + 2^{-n}.$$

Note that $x_u \neq x_v (u \neq v)$ means that there exists $a \in \{1, 2, \cdots, m\}$, s.t. $y_{1u} = y_{1v}, \cdots, y_{a-1,u} = y_{a-1,v}, y_{au} \neq y_{av}$, which means $OP_{f_{ia}}u$ and $OP_{f_{ia}}v$ are independent and random from $1'.2)$ of ψ because f_{i_a} is a random function, which means that $IP_{f_{j_b}}u$ and $IP_{f_{j_b}}v$ are random and independent for every $b \in \{1, 2, \cdots, m\}$ except for the case of the probability equaling or being smaller than $\delta(n)$, because $IP_{f_{j_b}}$ is $\delta(n)$-random from $1'.3)$ in $\delta(n)$-condition $1'$.

The total number of terms on the right side of (1) is

$$m \left(\begin{array}{c} k \\ 2 \end{array} \right) = \frac{mk(k-1)}{2} \leq \frac{mk^2}{2}.$$

Lemma 1 follows.

4.2 Pseudorandomness of SPN-structures and E2-like transformations

For applying $\delta(n)$-condition 1 to the SPN-structures and E2-like transformations, we introduce the next matrix operation.

Definition 10 *Let m be a positive integer. For the $m \times m$ matrix over $GF(2)$ $A = (a_{ij}), B = (b_{ij}))$, we define binary operation $\star$ by $A \star B = (\vee_{k=0}^{m}(a_{ik} \wedge b_{kj}))$, where $\vee$ represents binary operation "or", and $\wedge$ represents binary operation "and".*

Using this operation, we obtain the conditions under which the 3-round SPN-structure to yields PRF.

Lemma 2 *The 3-round SPN-structure with linear transformations A, B satisfies $1/2^n$-condition $1'$ if all components of $A \star B$ are 1.*

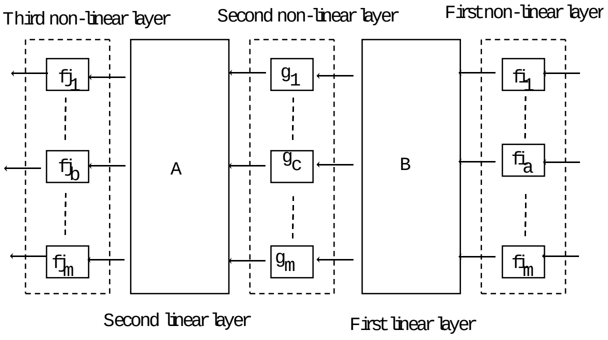

Figure 5: Pseudorandomness of 3-round SPN-structure

Proof. Let $A = (a_{ij}), B = (b_{ij})$ and $A \star B = (c_{ij})$, and we select $f_{i1}, \cdots, f_{i_m}, f_{j1}, \cdots, f_{j_m}$ as in Fig.5.

Then the condition that $IP_{f_{j_b}}$ depends on $OP_{f_{i_a}}$, is equivalent to $c_{ba} = 1$, because, from the independence of $g_1, \cdots, g_m$, this condition is equivalent to the condition that there exists g_c such that $IP_{f_{j_b}}$ depends on OP_{g_c} and IP_{g_c} depends on $OP_{f_{ia}}$

Next we prove that the randomness of OP_{fi_b} implies $\delta(n)$-randomness of $IP_{f_{j_b}}$ when $IP_{f_{j_b}}$ depends on $OP_{f_{i_a}}$. If $OP_{f_{ia}}, OP'_{f_{ia}}$ is random, the probability that a value on IP_{g_c} and a value on IP'_{g_c} are the same is $1/2^n$. This implies $IP_{f_{j_b}}$ is $1/2^n$-random because of the randomness of g_c and the independence of $g_1, \cdots, g_m$. The lemma was proved.

This lemma can be easily generalized for the case of n-round SPN-structures as follows.

Lemma 3 *The s-round SPN-structure with linear transformations $A_1, A_2, \cdots, A_{s-1}$ satisfies $1/2^n$-condition $1'$ if all components of $A_1 \star A_2 \star \cdots \star A_{s-1}$ are* 1.

Note. For the matrix P used as the linear transformation layer as in [K98], all components of $P \star P$ are 1. This implies it can yield PRFG (PRPG) with 3-rounds i.e. with only 24 cryptographic functions, whereas MISTY(1)-like transformations need 45(36) cryptographic functions and recursive Feistel structures need 27 cryptographic functions. This implies that, in this case, the SPN-structures is more secure than (recursive) Feistel structures in the viewpoint of pseudorandomness.

From this fact and lemma 1, we obtain the following theorem in the same way as Theorem 1 of [LR86],

Theorem 1 *Let $\mathcal{F}_{ij}$, for $1 \leq i \leq 3, 1 \leq j \leq 8$ be $3 \cdot 8$ independent (n, n, k, ϵ_{ij}) PRFs. Then the 3-round SPN-structures with linear transformation of E2,* $\mathrm{SPN}_{3,8}(\mathcal{F}_{11}, \cdots, \mathcal{F}_{38})$, *is a $(8n, 8n, k, \epsilon)$ PRF where $\epsilon = 8k^2 2^{-n} + \sum_{1 \leq i \leq 3} \sum_{1 \leq j \leq 8} \epsilon_{ij}$*

This theorem indicates that the 3-round SPN-structures with linear transformation of E2 are PRFG.

By the same argument, we can prove the next lemma.

Lemma 4 *The* 4*-round E2-like transformation satisfies* $1/2^n$*-condition 1'*

Proof. This can be proved by regarding the first nonlinear layer in the first s-box as $f_{i_1}, \cdots, f_{i_m}$, first nonlinear layer in the second s-box as $f_{i_{m+1}}, \cdots, f_{i_{2m}}$, second nonlinear layer in the last s-box and as $f_{j_1}, \cdots, f_{j_m}$, second nonlinear layer in the 4-th s-box as $f_{j_{m+1}}, \cdots, f_{j_{2m}}$, as shown in Fig.6.

From this lemma and lemma 1, we obtain the following theorem in the same way as Theorem 1 of [LR86],

Theorem 2 *Let* $\mathcal{F}_{ijk}$*, for* $1 \leq i \leq 4, 1 \leq j \leq 2, 1 \leq k \leq 8$ *be* $4 \cdot 2 \cdot 8$ *independent* $(n, n, k, \epsilon_{ijk})$ *PRFs. Then the* 4*-round E2-like transformations* $\mathrm{E2}_{4,2,8}(\mathcal{F}_{111}, \cdots, \mathcal{F}_{428})$ *is a* $(16n, 16n, k, \epsilon)$ *PRF where* $\epsilon = 16k^2 2^{-n} + \sum_{1 \leq i \leq 4} \sum_{1 \leq j \leq 2} \sum_{1 \leq k \leq 8} \epsilon_{ijk}$

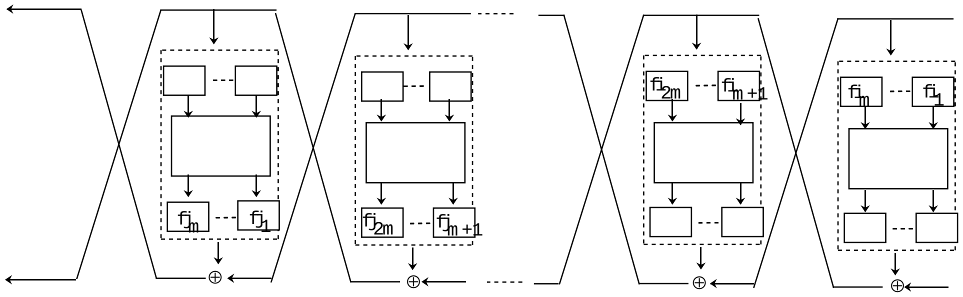

Figure 6: Pseudorandomness of E2-like transformation

This Theorem indicates that 4-round E2-like transformation are PRFG.

5 Maximum Average of Differential Probability of SPN-Structures

In this section, we consider the maximum average of differential probability of SPN-structures.

First we define the differentials of block ciphers. We consider the encryption of a pair of distinct plaintexts by an r-round iterated cipher. Here 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 plain-texts (or two cipher texts) X and X^* be defined as

$$\Delta X = X \oplus X^*.$$

From the pair of encryptions, one obtains the sequence of differences $\Delta X(0), \Delta X(1), \cdots, \Delta X(r)$ where $X(0) = X$ and $X(0)^* = X^*$ denote the plaintext pair (so that $\Delta X(0) = \Delta X$) and where $X(i)$ and $X^*(i)$ 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 $Z^{(i)}$.

Next we define the i-th round differential and maximum average of differential probabilities.

Definition 11 (LM92) *An i-round differential is the couple (α, β), where α is the differential of a pair of distinct plaintexts X and X^* and β is a possible difference for the resulting i-th round outputs $X(i)$ and $X^*(i)$. The probability of an i-round differential (α, β) is the conditional probability that β is the difference $\Delta X(i)$ of the cipher text pair after i rounds given that the plaintext pair (X, X^*) has difference $\Delta X = \alpha$ when the plaintext X and the subkeys $Z^{(1)}, \cdots, Z^{(i)}$ are independent and uniformly random. We denote this differential probability by $P(\Delta X(i) = \beta | \Delta X = \alpha)$.*

The probability of an s-round differential are known to be satisfying the following properties.

Lemma 5 (NK94) *The probability of an s-round differential equals*

$$P(\Delta X(s) = \beta(s) | \Delta X(0) = \beta(0)) =$$
$$\sum_{\beta(1)} \sum_{\beta(2)} \cdots \sum_{\beta(s-1)} \prod_{i=1}^{s} P(\Delta X(i) = \beta(i) | \Delta X(i-1) = \beta(i-1)).$$

We define the maximum average of differential probability as follows. This value is known to be the best measure to ensure that the block ciphers are secure against the differential attacks of block ciphers.

Definition 12 (NK94) *We define the maximum average of differential probability $ADP^{(s)}_{\max}$ by*

$$\mathrm{ADP}^{(s)}_{\max} = \max{}_{\alpha \neq 0, \beta} P(\Delta X(i) = \beta | \Delta X = \alpha).$$

Here we evaluate the maximum average of the differential probability in the case of the SPN-structure, where we assume all random functions are bijective. This value was considered to be too hard to evaluate in [K98], so they used the another approximate measure to estimate the security against differential attack. However, the following procedures suggest that this is easy to evaluate.

First we define the function ch : $\{GF(2)^n\}^m \to GF(2)^m, (x_1, \cdots, x_m) \longmapsto (y_1, \cdots, y_m)$ by

$$y_i = \begin{cases} 0 & \text{if } x_i = 0 \\ 1 & \text{otherwise,} \end{cases}$$

and we define the function $\mathrm{N}(P, \gamma, \delta)$ for $m \times m$ matrix P and $\gamma, \delta \in GF(2)^m$ by

$$\mathrm{N}(P, \gamma, \delta) = \#\{(\Delta X, \Delta Y) | \Delta Y = P \Delta X, \mathrm{ch}(\Delta X) = \gamma, \mathrm{ch}(\Delta Y) = \delta\}.$$

The procedure of calculating the maximum average of the differential probability in the case of the SPN-structure is as follows.

1) calculate $N(P, \gamma, \delta)$ for every $\gamma, \delta \in GF(2)^n$. For this calculation we define semi-order $\prec$ in $GF(2)^{2m}$ as follows.

$$a \prec b \Leftrightarrow (\forall i; (a(i) = 0 \Rightarrow b(i) = 0)) \wedge (a \neq b)$$

where we denote $a(i), b(i)$ as the i-th significant bit of a, b, respectively. We define

$$\mathrm{M}(P, \gamma, \delta) = \#\{(\Delta X, \Delta Y) | \Delta Y = P\Delta X, \mathrm{ch}(\Delta X) \preceq \gamma, \mathrm{ch}(\Delta Y) \preceq \delta\},$$

$\mathrm{M}(P, \gamma, \delta)$ can be easily calculated by simple rank calculation as follows.

$$\mathrm{M}(P, \gamma, \delta) = 2^{2m - \mathrm{rank}\left(\begin{pmatrix} A \quad E \\ F(\gamma, \delta) \end{pmatrix}\right)} - 1$$

where $F(\gamma, \delta)$ denotes the diagonal matrix whose (i, i) component equals the i-th significant bit of γ for $i = 1, \cdots, m$, or the $(i - m)$-th significant bit of δ for $i = m + 1, \cdots, 2m$.

$\mathrm{N}(P, \gamma, \delta)$ can be calculated recursively, using the following relations.

$$\mathrm{N}(P, \gamma, \delta) = \mathrm{M}(P, \gamma, \delta) - \sum_{\gamma' \prec \gamma, \delta' \prec \delta} N(P, \gamma', \delta')$$

2) calculate

$$P_1(\beta'(1), \beta'(0)) = \max_{\substack{\beta(0), \beta(1), \\ \mathrm{ch}(\beta(0)) = \beta'(0), \\ \mathrm{ch}(\beta(1)) = \beta'(1)}} P(\Delta X(1) = \beta(1) | \Delta X(0) = \beta(0))$$

for every $\beta'(1), \beta'(0) \in GF(2)^m$.

3) utilizing $N(P, \mathrm{ch}(\beta(i)), \mathrm{ch}(\beta(i-1)))$, calculate $P_i(\beta'(i), \beta'(i-1))$ recursively for every $\beta'(i), \beta'(i-1) \in GF(2)^m$.

$$\begin{aligned} & P_i(\beta'(i), \beta'(0)) = \\ & \sum_{\beta(i-1)} \max_{\substack{\beta^*(i-1), \beta(i), \\ \mathrm{ch}(\beta^*(i-1)) = \mathrm{ch}(\beta(i-1)), \\ \mathrm{ch}(\beta(i)) = \beta'(i)}} P(\Delta X(i) = \beta(i) | \Delta X(i-1) = \beta^*(i-1)) \\ & \qquad * P_{i-1}(\mathrm{ch}(\beta(i-1)), \beta'(0)) \\ = \quad & (2^n - 1) * \\ & \sum_{\beta'(i-1)} \max_{\substack{\beta(i-1), \beta(i), \\ \mathrm{ch}(\beta(i-1)) = \beta'(i-1), \\ \mathrm{ch}(\beta(i)) = \beta'(i)}} P(\Delta X(i) = \beta(i) | \Delta X(i-1) = \beta(i-1)) \\ & \qquad * P_{i-1}(\beta'(i-1), \beta'(0)) \\ = \quad & (2^n - 1) * \\ & \sum_{\beta'(i-1)} \mathrm{N}(P, \beta'(i), \beta'(i-1)) * p_{\max}^{h(\beta'(i))} * \ P_{i-1}(\beta'(i-1), \beta'(0)), \end{aligned}$$

where $p_{\max}$ is the maximum average of the differential probability of n-bit bijective s-boxes composing the non-linear layer.

By this procedure, we can exactly evaluate the maximum average of the differential probability under the assumption $P(\Delta X(i) = \beta(i)|\mathrm{ch}(\Delta X(i)) = \mathrm{ch}(\beta(i))) = 1/2^n$ for any $\beta(i)$ (We accept that is “impossible”, but the above has some validity as an ideal model.).

In the case of $m = 8$, we get

$$P(\Delta X(i) = \beta(i)|\Delta X(0) = \beta(0))$$
$$\leq \begin{cases} 255p^5(\text{for 2-round}), \\ 254p^7 + 255p^8 + p^9(\text{for 3-round}), \\ p^8 + 241p^9 + 284p^{10} + 162p^{11} + 206p^{12} + 230p^{13} + 214p^{14} + 108p^{15} \\ +222p^{16} + 73p^{17} + 193p^{18} + 206p^{19}(\text{for 4-round}), \\ p^8 + 154p^9 + 217p^{10} + 25p^{11} + 240p^{12} + 113p^{13} + 185p^{14} + 77p^{15} \\ +77p^{16} + 7p^{17} + 34p^{18} + 56p^{19} + 34p^{20} + 109p^{21} + 233p^{22} + 113p^{23} \\ +175p^{24} + 25p^{25} + 171p^{26} + 226p^{27} + 121p^{28} + 89p^{29} + 87p^{30} \\ +19p^{31} + 71p^{32} + 247p^{33}(\text{for 5-round}), \end{cases}$$

by the computer, and this indicates

$$P(\Delta X(i) = \beta(i)|\Delta X(0) = \beta(0)) \leq 1/2^{n-1}$$

for $m(\geq 4)$-round SPN-structures. This upper-bound is smaller than twice the maximum average of the differential probability of the functions constructing the nonlinear layer.

Without the assumption, the estimation is not so effective, but this can exactly evaluate the number of active s-boxes for all multiple passes. The evaluation in this section suggests that SPN-structure is a good structure in the viewpoint of immunity for differential attacks, even if we consider the multiple paths.

6 Conclusion

This paper examined the pseudorandomness of SPN-structures and E2-like transformations and showed that this characteristic can be easily calculated by some matrix calculation. Moreover, we examined the maximum average of the differential probability of the SPN-structure, and showed that this can be calculated recursively by a simple calculation. In AES candidate, SPN-structure is used in CRYPTON, E2, LOKI97, MARS, RC6, RIJNDAEL, SAFER, SERPENT. We conclude SPN-structure is better in pseudorandomness, a little weaker (but sufficiently strong) in immunity for differential attacks than (recursive) Feistel structures.

References

[BKR94] M. Bellare, J. Kilian and P. Rogaway, “The security of cipher block chaining.” Advances in Cryptology - Crypto 94 Proceedings, Lecture Notes in Computer Science Vol. 963, D. Coppersmith ed., Springer Verlag, 1994.

[GGM86] O. Goldreich, S. Goldwasser and S. Micali, “How to construct random functions,” Journal of the ACM, Vol. 33, No. 4, 1986, pp. 210-299.

[K98] M. Kanda et al. “A New 128-bit Block Cipher E2” Technical Report of IEICE. ISEC98-12.

[LM92] X. Lai, J. L. Massey and S. Murphy, “Markov Ciphers and Differential Cryptanalysis,” Advances in Cryptography-EUROCRYPTO ’91. Lecture Notes in Computer Science, Vol. 576. Springer-Verlag, Berlin, 1992, pages. 86-100.

[LR86] M. Luby and C. Rackoff, “How to construct pseudorandom permutations from pseudorandom functions,” STOC’86(also in SIAM-COMP.1988).

[M92] Ueli, M. Maurer, “A simplified and generalized treatment of Luby-Rackoff pseudorandom permutation generators,” In R. A. Rueppel, editor, Advances in Cryptology – EUROCRYPT 92, volume 658 of Lecture Notes in Computer Science, pages 239-255. Springer-Verlag, 24-28 May 1992.

[Ma96] M. Matsui, “New structure of block ciphers with provable security against differential and linear cryptanalysis,” In Dieter Grollman, editor, Fast Software Encryption: Third International Workshop, volume 1039 of Lecture Notes in Computer Science, pages 205-218, Cambridge, UK, 21-23 February 1996. Springer-Verlag.

[Ma97] M. Matsui, New block encryption algorithm MISTY. In Eli Biham, editor, Fast Software Encryption: 4th International Workshop, volume 1267 of Lecture Notes in Computer Science, pages 54-68, Haifa, Israel, 20-22 January 1997. Springer-Verlag

[NK94] K. Nyberg and L. R. Knudsen, “Provable security against a differential attack,” in Advances in Cryptology - EUROCRYTO’93, LNCS 765, pages 55-64, Springer-Verlag, Berlin, 1994.

[SZ96] K. Sakurai and Y. Zheng, “On Pseudo Randomness from Block Ciphers,” SCIS96.

[S97] M. Sugita, “Pseudorandomness of a Block Cipher with Recursive Structures.” Technical Report of IEICE. ISEC97-9.

[S97-2] M. Sugita, “Pseudorandomness of Block Cipher MISTY1.” Technical Report of IEICE. ISEC97-19.

[Sh49] C. Shannon, “Communication theory of Secrecy system.” Bell Systems Technical Journal, 28(4), 656-715 (1949).