Provable Security of KASUMI and 3GPP Encryption Mode $f8$

Ju-Sung Kang[1], Sang-Uk Shin[1], Dowon Hong[1], and Okyeon Yi[2]

[1] Section 0741, Information Security Technology Division, ETRI
161 Kajong-Dong, Yusong-Gu, Taejon, 305-350, KOREA
{jskang,shinsu,dwhong}@etri.re.kr
[2] Department of Mathematics, Kookmin University
Jeongreung3-Dong, Seongbuk-Gu, Seoul, 136-702, KOREA
oyyi@kmu.kookmin.ac.kr

Abstract. Within the security architecture of the 3GPP system there is a standardised encryption mode $f8$ based on the block cipher KASUMI. In this work we examine the pseudorandomness of the block cipher KASUMI and the provable security of $f8$. First we show that the three round KASUMI is not a pseudorandom permutation ensemble but the four round KASUMI is a pseudorandom permutation ensemble under the adaptive distinguisher model by investigating the properties of the round functions in a clear way. Second we provide the upper bound on the security of $f8$ mode under the reasonable assumption from the first result by means of the left-or-right security notion.

1 Introduction

There is a standardised encryption algorithm $f8$ within the security architecture of the 3GPP(3rd Generation Partnership Project) system and this algorithm is based on the block cipher KASUMI that produces a 64-bit output from a 64-bit input under the control of an 128-bit key[12]. To guarantee the message confidentiality over a radio access link of W-CDMA IMT-2000, $f8$ encryption mode with KASUMI has been proposed. The purpose of this work is to investigate the pseudorandomness of the block cipher KASUMI and the provable security of $f8$.

A block cipher can be regarded as a family of permutations on a message space indexed by a secret key. Luby-Rackoff[7] introduced a theoretical model for the security of block ciphers by using the notion of pseudorandom and super-pseudorandom permutations. A pseudorandom permutation can be interpreted as a block cipher that no attacker with polynomially many encryption queries can distinguish between the block cipher and the perfect random permutation. In [7], Luby and Rackoff used the DES-type transformation in order to construct a pseudorandom permutation from a pseudorandom function. They showed that the DES-type transformation with three rounds yielded $2n$-bit pseudorandom permutation under the assumption that each round function was an n-bit pseudorandom function. Sakurai-Zheng[11] showed that the three round MISTY-type transformation was not a pseudorandom permutation ensemble. MISTY-type

C. Boyd (Ed.): ASIACRYPT 2001, LNCS 2248, pp. 255–271, 2001.

transformation[8,9] was another two-block structure different from DES-type. Recently, Gilbert-Minier[4] and Kang et al.[6] showed independently that the four round MISTY-type transformation was a pseudorandom permutation.

The overall structure of KASUMI is the DES-type, but its round function FO composed of three round MISTY-type transformation which is not a pseudorandom function. Thus we cannot straightforwardly apply the Luby-Rackoff's result to KASUMI. FO function within KASUMI has FI function as its component function which is composed of four round unbalanced MISTY-type transformation. We show that this is a pseudorandom permutation. And we prove that the three round KASUMI is not a pseudorandom permutation but the four round KASUMI is a pseudorandom permutation. In [6], the authors investigated the pseudorandomness of KASUMI for non-adaptive distinguishers. In this paper we consider the security model for adaptive distinguishers similar to the approach of Naor and Reingold[10] and investigate the properties of the round function of KASUMI more precisely than the previous results like [4] and [6].

On the other hand $f8$ is one of the modes of operation for block ciphers. Several modes of operation for block ciphers have been proposed to encrypt plaintext blocks more than one block and to fulfil varying application requirements. As standardized modes of operation, ECB(electronic codebook), CBC(cipher block chaining), CFB(cipher feedback) and OFB(output feedback) are known[3]. 3GPP $f8$ encryption mode can be seen as a variant of OFB mode.

Proving the security of modes of operation started by Bellare et al.[1] in 1994 who analyzed the security of CBC MAC mode. In 1997, Bellare et al.[2] introduced the security notions of the symmetric encryption scheme and proved the security of CTR mode and CBC mode. Recently, Alkassar et al.[13] analyzed the security of CFB mode and proposed the OCFB mode which improved the performance of CFB mode. In this paper we show that 3GPP $f8$ encryption mode is secure by means of the left-or-right security notion. To prove this fact we should have the assumption that the underlying block cipher KASUMI is secure. This assumption is reasonable since by the first our result we already obtain that KASUMI is a pseudorandom permutation ensemble.

2 Pseudorandomness of the Block Cipher KASUMI

2.1 Preliminaries

Let I_n denote the set of all n-bit strings and $\mathcal{P}_n$ be the set of all permutations from I_n to itself where n is a positive integer. That is, $\mathcal{P}_n = \{\pi : I_n \rightarrow I_n \mid \pi \text{ is a bijection}\}$. We define an n-bit perfect random permutation as an uniformly drawn element of $\mathcal{P}_n$.

Definition 1. *$\mathcal{P}_n$ is called the UPE(uniform permutation ensemble) if all permutations in $\mathcal{P}_n$ are uniformly distributed. That is, for any permutation $\pi \in \mathcal{P}_n$, $Pr(\pi) = \frac{1}{2^n!}$.*

We consider the following security model. Let $\mathcal{D}$ be a computationally unbounded distinguisher with an oracle $\mathcal{O}$. The oracle $\mathcal{O}$ chooses randomly a permutation π from the UPE $\mathcal{P}_n$ or from a permutation ensemble $\Lambda_n \subset \mathcal{P}_n$. For

an n-bit block cipher, Λ_n is the set of permutations determined by all the secret keys. The purpose of the distinguisher $\mathcal{D}$ is to distinguish whether the oracle $\mathcal{O}$ implements the UPE $\mathcal{P}_n$ or Λ_n.

Definition 2. *Let $\mathcal{D}$ be a distinguisher, $\mathcal{P}_n$ be the UPE, and Λ_n be a permutation ensemble obtained from a block cipher. Then the advantage $ADV_{\mathcal{D}}$ of $\mathcal{D}$ is defined by*

$$ADV_{\mathcal{D}} = |Pr(\mathcal{D} \text{ outputs } 1 \mid \mathcal{O} \leftarrow \mathcal{P}_n) - Pr(\mathcal{D} \text{ outputs } 1 \mid \mathcal{O} \leftarrow \Lambda_n)| \ ,$$

where $\mathcal{O} \leftarrow \mathcal{P}_n$ and $\mathcal{O} \leftarrow \Lambda_n$ denote that $\mathcal{O}$ implements $\mathcal{P}_n$ and Λ_n, respectively.

Assume that the distinguisher $\mathcal{D}$ is restricted to make at most $poly(n)$ queries to the oracle $\mathcal{O}$, where $poly(n)$ is some polynomial in n. We call $\mathcal{D}$ is a pseudorandom distinguisher if it queries x and the oracle answers $y = \pi(x)$, where π is a randomly chosen permutation by $\mathcal{O}$. We say that $\mathcal{D}$ is a super-pseudorandom distinguisher if it is a pseudorandom distinguisher and also makes a query y and receives $x = \pi^{-1}(y)$ from the oracle $\mathcal{O}$.

Definition 3. *A function $h : \mathbb{N} \rightarrow \mathbb{R}$ is called negligible if for any constant $c > 0$ and all sufficiently large $n \in \mathbb{N}$, $h(n) < \frac{1}{n^c}$.*

Definition 4. *Let Λ_n be an efficiently computable permutation ensemble. Then Λ_n is called a PPE(pseudorandom permutation ensemble) if $ADV_{\mathcal{D}}$ is negligible for any pseudorandom distinguisher $\mathcal{D}$.*

Definition 5. *Let Λ_n be an efficiently computable permutation ensemble. Then we call Λ_n is a SPPE(super-pseudorandom permutation ensemble) if $ADV_{\mathcal{D}}$ is negligible for any super-pseudorandom distinguisher $\mathcal{D}$.*

In Definition 4 and 5, a permutation ensemble is efficiently computable if all permutations in the ensemble can be computed efficiently. See [10] for the rigorous definition of this. It is reasonable assumption that Λ_n is an efficiently computable permutation ensemble if it is obtained from an n-bit block cipher. Hence we assume that any permutation ensemble obtained from a block cipher is efficiently computable.

We define two transformations, DES-type and MISTY-type, which are obtained from two representative structures of current block ciphers. Let $\mathcal{F}_n$ denote the set of all functions from I_n to itself. We call briefly f is an n-bit function(resp. permutation) where $f \in \mathcal{F}_n$(resp. $f \in \mathcal{P}_n$).

Definition 6. *For any n-bit function $f \in \mathcal{F}_n$, $2n$-bit DES-type permutation $\mathbf{D}_f \in \mathcal{P}_{2n}$ is defined by $\mathbf{D}_f(L, R) = (R, L \oplus f(R))$, where $L, R \in I_n$.*

Definition 7. *For any n-bit permutation $f \in \mathcal{P}_n$, $2n$-bit MISTY-type permutation $\mathbf{M}_f \in \mathcal{P}_{2n}$ is defined by $\mathbf{M}_f(L, R) = (R, f(L) \oplus R)$, where $L, R \in I_n$.*

Several noticeable results about the pseudorandomness of DES-type and MISTY-type transformations are as follows. It is aware that PFE(pseudorandom function ensemble) can be similarly defined as Definition 4 by considering function space instead of permutation space.

- $\mathbf{D}_{f_2} \circ \mathbf{D}_{f_1}$ is not a $2n$-bit PPE and $\mathbf{D}_{f_3} \circ \mathbf{D}_{f_2} \circ \mathbf{D}_{f_1}$ is not a $2n$-bit SPPE, although all f_i's$(i = 1, 2, 3)$ are independently chosen from an n-bit PFE[7].
- $\mathbf{D}_{f_3} \circ \mathbf{D}_{f_2} \circ \mathbf{D}_{f_1}$ is a $2n$-bit PPE and $\mathbf{D}_{f_4} \circ \mathbf{D}_{f_3} \circ \mathbf{D}_{f_2} \circ \mathbf{D}_{f_1}$ is a $2n$-bit SPPE if all f_i's$(i = 1, 2, 3, 4)$ are independently chosen from an n-bit PFE[7].
- $\mathbf{M}_{f_3} \circ \mathbf{M}_{f_2} \circ \mathbf{M}_{f_1}$ is not a $2n$-bit PPE and $\mathbf{M}_{f_4} \circ \mathbf{M}_{f_3} \circ \mathbf{M}_{f_2} \circ \mathbf{M}_{f_1}$ is not a $2n$-bit SPPE, although each $f_i (i = 1, 2, 3, 4)$ is chosen independently from an n-bit PPE[4,11].
- $\mathbf{M}_{f_4} \circ \mathbf{M}_{f_3} \circ \mathbf{M}_{f_2} \circ \mathbf{M}_{f_1}$ is a $2n$-bit PPE and $\mathbf{M}_{f_5} \circ \mathbf{M}_{f_4} \circ \mathbf{M}_{f_3} \circ \mathbf{M}_{f_2} \circ \mathbf{M}_{f_1}$ is a $2n$-bit SPPE, where all f_i's$(i = 1, 2, 3, 4, 5)$ are independently chosen from an n-bit PPE[4,5,6].

On the other hand KASUMI is a modified version of the block cipher MISTY1[9] and we can classify the permutation of KASUMI into the following three stages:

- The overall permutation of KASUMI is a 64-bit permutation composed of the eight round DES-type permutation with the two round permutation FO and FL.
- FO function is a 32-bit permutation composed of the three round MISTY-type transformation with the round permutation FI.
- FI function is a 16-bit permutation which is composed of the four round unbalanced MISTY-type transformation obtained from 7-bit S-box $S7$ and 9-bit S-box $S9$.

First we show that FI function is a 16-bit PPE by examining the pseudorandomness of unbalanced MISTY-type transformation. Second we prove that three round KASUMI is not a 64-bit PPE but four round KASUMI is a 64-bit PPE on the base of the first result. Note that FO function is not a 32-bit PPE, so it doesn't seem that the three round DES-type permutation of KASUMI is a 64-bit PPE as the Luby-Rackoff cipher. Since the FL function is to round key mixing, we can omit FL function in order to analyze the pseudoranomness of KASUMI.

2.2 Pseudorandomness of the Unbalanced MISTY-Type Transformation

We describe simple but useful two lemmas which their proofs are given in [6].

Lemma 1. *Let π be a permutation chosen from the UPE $\mathcal{P}_n$. Then for any $x_1 \neq x_2, y \in I_n$,*

$$Pr(\pi(x_1) \oplus \pi(x_2) = y) = \begin{cases} \frac{1}{2^n - 1} & \text{if } y \neq 0\ , \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 2. *Let π_1 and π_2 be two permutations independently chosen from the UPE $\mathcal{P}_n$. Then for any $a, b, c, d, y \in I_n$,*

$$Pr\left(\pi_1(a) \oplus \pi_1(b) \oplus \pi_2(c) \oplus \pi_2(d) = y\right) < \frac{1}{2^{n-1}}\ ,\ \text{for } n \geq 2.$$

Now we define two unbalanced MISTY-type transformations to examine accurately the pseudorandomness of FI function.

Definition 8. *Let n and m be two positive integers such that $m \leq n$. Then for any n-bit permutation f and m-bit permutation g, two $(n+m)$-bit unbalanced MISTY-type transformations $\overline{\mathbf{M}}_f \in \mathcal{P}_{n+m}$ and $\widehat{\mathbf{M}}_g \in \mathcal{P}_{n+m}$ are defined by*

$$\overline{\mathbf{M}}_f(L,R) = (R, f(L) \oplus \overline{R}) \in I_m \times I_n \ , \ \ \forall (L,R) \in I_n \times I_m$$

and

$$\widehat{\mathbf{M}}_g(L,R) = (R, g(L) \oplus \widehat{R}) \in I_n \times I_m \ , \ \ \forall (L,R) \in I_m \times I_n \ ,$$

where for any n-bit vector x, $\widehat{x}$ denotes the m-bit value obtained by discarding the $n-m$ most-significant end and for any m-bit vector y, $\overline{y}$ denotes the n-bit value obtained by adding $n-m$ zero bits to the most-significant end.

Note that the FI function of KASUMI can be represented as 16-bit permutation $\widehat{\mathbf{M}}_{f_4} \circ \overline{\mathbf{M}}_{f_3} \circ \widehat{\mathbf{M}}_{f_2} \circ \overline{\mathbf{M}}_{f_1}$, where f_1, f_3 are 9-bit permutations and f_2, f_4 are 7-bit permutations. The pseudorandomness of the FI function is guaranteed by the following theorem.

Theorem 1. *Let for any positive integer n and m such that $m \leq n$, $f_1, f_3 \in \mathcal{P}_n$ and $f_2, f_4 \in \mathcal{P}_m$ be independently chosen from two n-bit and m-bit PPEs, respectively. Then the four round unbalanced MISTY-type transformation $\widehat{\mathbf{M}}_{f_4} \circ \overline{\mathbf{M}}_{f_3} \circ \widehat{\mathbf{M}}_{f_2} \circ \overline{\mathbf{M}}_{f_1}$ is an $(n+m)$-bit PPE.*

Recall that a pseudorandom distinguisher $\mathcal{D}$ can make query x and the oracle $\mathcal{O}$ answers $y = \pi(x)$, where π is a randomly chosen permutation by $\mathcal{O}$. Now we assume that $\mathcal{D}$ makes exactly q queries and refer to the sequence $\{(x^{(1)}, y^{(1)}), \cdots, (x^{(q)}, y^{(q)})\}$ of all query-answer pairs as the $\mathcal{D}$-transcript, where $q = poly(n)$. We consider an adaptive pseudorandom distinguisher as the following definition.

Definition 9. *$\mathcal{D}$ is called an adaptive pseudorandom distinguisher if it has a transcript $\{(x^{(1)}, y^{(1)}), \cdots, (x^{(q)}, y^{(q)})\}$ and a function $C_{\mathcal{D}}$ of $\mathcal{D}$-transcript such that for every $2 \leq i \leq q$,*

$$x^{(i)} = C_{\mathcal{D}}(\{(x^{(1)}, y^{(1)}), \cdots, (x^{(i-1)}, y^{(i-1)})\})$$

and

$$\textit{the ouput of } \mathcal{D} = C_{\mathcal{D}}(\{(x^{(1)}, y^{(1)}), \cdots, (x^{(q)}, y^{(q)})\}) \ .$$

Under the adaptive distinguisher model, for any i-th query of $\mathcal{D}$ is fully determined by the first $i-1$ query-answer pairs and $\mathcal{D}$'s output is a function of its transcript. Throughout this paper we assume that all queries are distinct.

To prove the Theorem 1, we formally define a bad event and estimate its probability.

Definition 10. *For any n-bit permutation f_1 and m-bit permutation f_2, $BAD(f_1, f_2)$ is defined as the set of all $\mathcal{D}$-transcripts $\sigma = \{(x^{(1)}, y^{(1)}), \cdots, (x^{(q)}, y^{(q)})\}$ satisfying: $\exists 1 \le i < j \le q$ such that*

$$f_1(x_L^{(i)}) \oplus \overline{x_R^{(i)}} = f_1(x_L^{(j)}) \oplus \overline{x_R^{(j)}}$$

or

$$f_2(x_R^{(i)}) \oplus \widehat{f_1(x_L^{(i)})} \oplus x_R^{(i)} = f_2(x_R^{(j)}) \oplus \widehat{f_1(x_L^{(j)})} \oplus x_R^{(j)} \ ,$$

where $x^{(i)} = (x_L^{(i)}, x_R^{(i)}) \in I_n \times I_m$ for all $1 \le i \le q$.

Lemma 3. *Let f_1 and f_2 be chosen independently from UPE $\mathcal{P}_n$ and UPE $\mathcal{P}_m$, respectively. Then for any $\mathcal{D}$-transcript $\sigma = \{(x^{(1)}, y^{(1)}), \cdots, (x^{(q)}, y^{(q)})\}$ and $n \ge m \ge 2$,*

$$Pr(\sigma \in BAD(f_1, f_2)) < (q^2 - q)\left(\frac{1}{2^n} + \frac{1}{2^m}\right) \ .$$

Proof. By definition, $\sigma \in BAD(f_1, f_2)$ if there exist $1 \le i < j \le q$ such that either

$$f_1(x_L^{(i)}) \oplus \overline{x_R^{(i)}} = f_1(x_L^{(j)}) \oplus \overline{x_R^{(j)}}$$

or

$$f_2(x_R^{(i)}) \oplus \widehat{f_1(x_L^{(i)})} \oplus x_R^{(i)} = f_2(x_R^{(j)}) \oplus \widehat{f_1(x_L^{(j)})} \oplus x_R^{(j)} \ .$$

For any given i and j, we estimate probabilities of these two events. We have the following three cases.

Case 1: $x_L^{(i)} \ne x_L^{(j)}$ and $x_R^{(i)} = x_R^{(j)}$. Since f_1 is a permutation,

$$Pr\left(f_1(x_L^{(i)}) \oplus \overline{x_R^{(i)}} = f_1(x_L^{(j)}) \oplus \overline{x_R^{(j)}}\right) = Pr\left(f_1(x_L^{(i)}) = f_1(x_L^{(j)})\right) = 0 \ .$$

Observe that, by the similar result to Lemma 1

$$\begin{aligned} &Pr\left(f_2(x_R^{(i)}) \oplus \widehat{f_1(x_L^{(i)})} \oplus x_R^{(i)} = f_2(x_R^{(j)}) \oplus \widehat{f_1(x_L^{(j)})} \oplus x_R^{(j)}\right) \\ &= Pr\left(\widehat{f_1(x_L^{(i)})} = \widehat{f_1(x_L^{(j)})}\right) = 2^n \cdot \frac{2^{n-m} \cdot (2^n - 2)!}{2^n!} = \frac{2^{n-m}}{2^n - 1} \ . \end{aligned}$$

Case 2: $x_L^{(i)} = x_L^{(j)}$ and $x_R^{(i)} \ne x_R^{(j)}$. In this case the probability of the first event is equal to $Pr(\overline{x_R^{(i)}} = \overline{x_R^{(j)}}) = 0$. By Lemma 1, the probability of the second event is estimated as

$$Pr\left(f_2(x_R^{(i)}) \oplus f_2(x_R^{(j)}) = x_R^{(i)} \oplus x_R^{(j)}\right) = \frac{1}{2^m - 1} \ .$$

Case 3: $x_L^{(i)} \ne x_L^{(j)}$ and $x_R^{(i)} \ne x_R^{(j)}$. By Lemma 1, the probability of the first event is estimated as

$$Pr\left(f_1(x_L^{(i)}) \oplus f_1(x_L^{(j)}) = \overline{x_L^{(i)}} \oplus \overline{x_L^{(j)}}\right) = \frac{1}{2^n - 1} \ .$$

Similarly, by Lemma 2, the probability of the second event is also estimated as

$$Pr\left(\widehat{f_1(x_L^{(i)})} \oplus \widehat{f_1(x_L^{(j)})} \oplus f_2(x_R^{(i)}) \oplus f_2(x_R^{(j)}) = x_R^{(i)} \oplus x_R^{(j)}\right) < \frac{1}{2^m - 1} ,$$

since $n \geq m \geq 2$.

Hence, for any case, we obtain that

$$Pr\left(f_1(x_L^{(i)}) \oplus \overline{x_R^{(i)}} = f_1(x_L^{(j)}) \oplus \overline{x_R^{(j)}}\right) < \frac{1}{2^{n-1}}$$

and

$$Pr\left(f_2(x_R^{(i)}) \oplus \widehat{f_1(x_L^{(i)})} \oplus x_R^{(i)} = f_2(x_R^{(j)}) \oplus \widehat{f_1(x_L^{(j)})} \oplus x_R^{(j)}\right) < \frac{1}{2^{m-1}} .$$

Therefore

$$Pr\left(\sigma \in BAD(f_1, f_2)\right) < \binom{q}{2}\left(\frac{1}{2^{n-1}} + \frac{1}{2^{m-1}}\right) < (q^2 - q)\left(\frac{1}{2^n} + \frac{1}{2^m}\right) . \quad \square$$

Definition 11. *Let Λ_{n+m} be the $(n+m)$-bit permutation ensemble obtained from $\Lambda_{n+m}(f_1, f_2, f_3, f_4) = \widehat{\mathbf{M}}_{f_4} \circ \overline{\mathbf{M}}_{f_3} \circ \widehat{\mathbf{M}}_{f_2} \circ \overline{\mathbf{M}}_{f_1}$. Then $T_{\mathcal{P}_{n+m}}$ and $T_{\Lambda_{n+m}}$ are defined by the random variables such that $T_{\mathcal{P}_{n+m}}$ is the $\mathcal{D}$-transcript when the oracle $\mathcal{O}$ implements the UPE $\mathcal{P}_{n+m}$ and $T_{\Lambda_{n+m}}$ is the $\mathcal{D}$-transcript when the oracle $\mathcal{O}$ implements the permutation ensemble Λ_{n+m}.*

Lemma 4. *Let Λ_{n+m} be the $(n+m)$-bit permutation ensemble of all $\Lambda_{n+m}(f_1, f_2, f_3, f_4)$ such that $f_1, f_3 \in \mathcal{P}_n$ and $f_2, f_4 \in \mathcal{P}_m$ are independently chosen from the n-bit and m-bit UPEs, respectively. Then for any $\mathcal{D}$-transcript $\sigma = \{(x^{(1)}, y^{(1)}), \cdots, (x^{(q)}, y^{(q)})\}$,*

$$\left|Pr\left(T_{\Lambda_{n+m}} = \sigma \mid \ \sigma \notin BAD(f_1, f_2)\right) - Pr\left(T_{\mathcal{P}_{n+m}} = \sigma\right)\right| < \varepsilon_{n,m,q} ,$$

where

$$\varepsilon_{n,m,q} = \frac{1}{2^{n+m}(2^n - 1)(2^m - 1)\cdots(2^n - q + 1)(2^m - q + 1)} .$$

Proof. For any possible $\mathcal{D}$-transcript we have that

$$Pr\left(T_{\mathcal{P}_{n+m}} = \sigma\right) = \frac{(2^{n+m} - q)!}{2^{n+m}!} .$$

Consider any specific n-bit permutation f_1 and m-bit permutation f_2 such that $\sigma \notin BAD(f_1, f_2)$. Note that $T_{\Lambda_{n+m}} = \sigma$ if and only if for all $1 \leq i \leq q$, $y^{(i)} = \Lambda_{n+m}(x^{(i)})$. Since $\Lambda_{n+m} = \widehat{\mathbf{M}}_{f_4} \circ \overline{\mathbf{M}}_{f_3} \circ \widehat{\mathbf{M}}_{f_2} \circ \overline{\mathbf{M}}_{f_1}$,

$$y^{(i)} = \Lambda_{n+m}(x^{(i)}) \Leftrightarrow f_3(L_2^{(i)}) = y_L^{(i)} \oplus \overline{R_2^{(i)}} \in I_n \text{ and } f_4(R_2^{(i)}) = \widehat{y_L^{(i)}} \oplus y_R^{(i)} \in I_m ,$$

where $(L_2^{(i)}, R_2^{(i)}) = \widehat{\mathbf{M}}_{f_2} \circ \overline{\mathbf{M}}_{f_1}(x_L^{(i)}, x_R^{(i)})$. By definition of $BAD(f_1, f_2)$, if $\sigma \notin BAD(f_1, f_2)$, then $L_2^{(i)} \neq L_2^{(j)}$ and $R_2^{(i)} \neq R_2^{(j)}$ for all $1 \leq i \neq j \leq q$. Therefore, since f_3 and f_4 are independently chosen from the UPEs $\mathcal{P}_n$ and $\mathcal{P}_m$, respectively, we obtain that

$$Pr\left(T_{\Lambda_{n+m}} = \sigma \mid \ \sigma \notin BAD(f_1, f_2)\right) = \frac{(2^n - q)!}{2^n!} \cdot \frac{(2^m - q)!}{2^m!} \ ,$$

which complete the assertion. □

Proof of Theorem 1: It suffices to show the assertion under the assumption that $f_1, f_3 \in \mathcal{P}_n$ and $f_2, f_4 \in \mathcal{P}_m$ be independently chosen from two n-bit and m-bit UPEs, respectively. Let Λ_{n+m} be the $(n+m)$-bit permutation ensemble of all $\Lambda_{n+m}(f_1, f_2, f_3, f_4) = \widehat{\mathbf{M}}_{f_4} \circ \overline{\mathbf{M}}_{f_3} \circ \widehat{\mathbf{M}}_{f_2} \circ \overline{\mathbf{M}}_{f_1}$ and Θ be the set of all $\mathcal{D}$-transcripts σ such that the output of $\mathcal{D}$ is $C_{\mathcal{D}}(\sigma) = 1$. Then

$$\begin{aligned}
& ADV_{\mathcal{D}} \\
&= \left|Pr\left(C_{\mathcal{D}}(T_{\Lambda_{n+m}}) = 1\right) - Pr\left(C_{\mathcal{D}}(T_{\mathcal{P}_{n+m}}) = 1\right)\right| \\
&\leq \sum_{\sigma \in \Theta} Pr(\sigma \notin BAD(f_1, f_2)) \\
&\qquad \cdot \left|Pr\left(T_{\Lambda_{n+m}} = \sigma \mid \ \sigma \notin BAD(f_1, f_2)\right) - Pr\left(T_{\mathcal{P}_{n+m}} = \sigma\right)\right| \qquad (1) \\
&\quad + \sum_{\sigma \in \Theta} Pr\left(T_{\Lambda_{n+m}} = \sigma, \sigma \in BAD(f_1, f_2)\right) \qquad (2) \\
&\quad + \sum_{\sigma \in \Theta} Pr\left(\sigma \in BAD(f_1, f_2)\right) \cdot Pr(T_{\mathcal{P}_{n+m}} = \sigma). \qquad (3)
\end{aligned}$$

By Lemma 4, the term (1) is bounded above by $\varepsilon_{n,m,q}$ and by Lemma 3, the value of (3) is bounded by

$$\max_{\sigma \in \Theta} Pr(\sigma \in BAD(f_1, f_2)) \cdot Pr\left(\cup_{\sigma \in \Theta}\{T_{\mathcal{P}_{n+m}} = \sigma\}\right) < (q^2 - q)\left(\frac{1}{2^n} + \frac{1}{2^m}\right) \ .$$

On the other hand, by Lemma 3, the value of (2) is estimated as

$$\begin{aligned}
& \sum_{\sigma \in \Theta} Pr\left(T_{\Lambda_{n+m}} = \sigma, \sigma \in BAD(f_1, f_2)\right) \\
&= \sum_{\sigma \in \Theta} Pr(T_{\Lambda_{n+m}} = \sigma) \cdot Pr\left(\sigma \in BAD(f_1, f_2) \mid \ T_{\Lambda_{n+m}} = \sigma\right) \\
&< (q^2 - q)\left(\frac{1}{2^n} + \frac{1}{2^m}\right) \ .
\end{aligned}$$

Therefore we can conclude that

$$ADV_{\mathcal{D}} < 2(q^2 - q)\left(\frac{1}{2^n} + \frac{1}{2^m}\right) + \varepsilon_{n,m,q} \ ,$$

which is negligible. □

2.3 Pseudorandomness of KASUMI

From Theorem 1, it becomes a reasonable assumption that FI function of KASUMI is a PPE. In order to investigate the pseudorandomness of KASUMI, we use a simplified figure of KASUMI. The four round simplified KASUMI is illustrated in Figure 1, where $x = (x_1, x_2, x_3, x_4)$ denotes a $4n$-bit input value, $w = (w_1, w_2, w_3, w_4)$, $y = (y_1, y_2, y_3, y_4)$, and $z = (z_1, z_2, z_3, z_4)$ denote corresponding outputs of the two, three, and four round KASUMI, respectively. Each of x_i, w_i, y_i, and z_i is an n-bit value. By the following theorem, we obtain the fact that three round of KASUMI is insufficient to be a PPE.

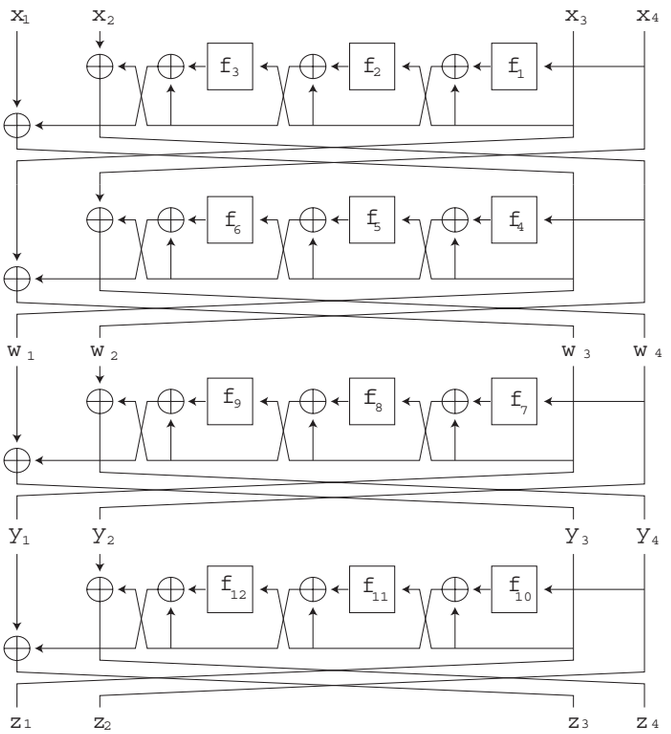

Fig. 1. Simplified four round KASUMI

Theorem 2. *The three round simplified KASUMI is not a $4n$-bit PPE though f_i's($i = 1, \cdots, 9$) of Figure 1 are independently chosen from an n-bit PPE.*

Proof. Let Λ_{4n} be the set of all permutations over I_{4n} obtained from the three round simplified KASUMI. Consider a distinguisher $\mathcal{D}$ such as follows:

1. $\mathcal{D}$ chooses four $4n$-bit queries $x^{(1)}$, $x^{(2)}$, $x^{(3)}$, and $x^{(4)}$ such that

$$x^{(1)} = (0, 0, x_3, x_4) \ , \quad x^{(2)} = (x_1, 0, x_3, x_4) \ ,$$

$$x^{(3)} = (0, x_2, x_3, x_4) \ , \quad x^{(4)} = (x_1, x_2, x_3, x_4) \ ,$$

   where $x_1 \neq 0 \neq x_2$ and x_3, x_4 are fixed n-bit values.
2. $\mathcal{D}$ sends these four queries to the oracle $\mathcal{O}$ and receives the corresponding answers $(y_1^{(i)}, y_2^{(i)}, y_3^{(i)}, y_4^{(i)})(i = 1, 2, 3, 4)$ from the oracle.
3. $\mathcal{D}$ outputs 1 if and only if

$$y_2^{(1)} \oplus y_2^{(2)} \oplus y_2^{(3)} \oplus y_2^{(4)} = 0 \ .$$

If the oracle implements the UPE $\mathcal{P}_{4n}$, then we obtain that

$$\begin{aligned} Pr(\mathcal{D} \text{ outputs } 1 \mid \mathcal{O} \leftarrow \mathcal{P}_{4n}) &\leq \frac{2^{4n}(2^{4n}-1)(2^{4n}-2)2^{3n}(2^{4n}-4)!}{2^{4n}!} \\ &= \frac{2^{3n}}{2^{4n}-3} \leq \frac{1}{2^{n-1}} \ . \end{aligned}$$

On the other hand, if $\mathcal{O}$ implements Λ_{4n}, then for $x^{(1)} = (0, 0, x_3, x_4)$, $x^{(2)} = (x_1, 0, x_3, x_4)$, $x^{(3)} = (0, x_2, x_3, x_4)$, and $x^{(4)} = (x_1, x_2, x_3, x_4)$, we can see from Figure 1 that the corresponding $2n$-bit inputs of the second round are

$$(F_1(x_3, x_4)|_L, F_1(x_3, x_4)|_R) \ , \quad (F_1(x_3, x_4)|_L, x_1 \oplus F_1(x_3, x_4)|_R) \ ,$$

$$(x_2 \oplus F_1(x_3, x_4)|_L, F_1(x_3, x_4)|_R) \ , \quad (x_1 \oplus F_1(x_3, x_4)|_L, x_2 \oplus F_1(x_3, x_4)|_R)$$

respectively, where $F_1 = \mathbf{M}_{f_3} \circ \mathbf{M}_{f_2} \circ \mathbf{M}_{f_1}$ and $(x|_L, x|_R)$ denote the left and right n-bit block of $2n$-bit value x. Thus we obtain by the similar argument of Sakurai-Zheng[11] that

$$y_2^{(1)} \oplus y_2^{(2)} \oplus y_2^{(3)} \oplus y_2^{(4)} = 0$$

with probability 1.

Consequently we obtain that

$$\begin{aligned} ADV_{\mathcal{D}} &= |Pr(\mathcal{D} \text{ outputs } 1 \mid \mathcal{O} \leftarrow \mathcal{P}_{4n}) - Pr(\mathcal{D} \text{ outputs } 1 \mid \mathcal{O} \leftarrow \Lambda_{4n})| \\ &\geq 1 - \frac{1}{2^{n-1}} \ , \end{aligned}$$

which is non-negligible. □

The following theorem guarantees that the four or more round KASUMI is a pseudorandom permutation ensemble.

Theorem 3. *If f_i's($i = 1, 2, \cdots, 12$) in Figure 1 are independently chosen from an n-bit PPE, then the four round KASUMI is a $4n$-bit PPE.*

From Figure 1, we can see that the second round output w_3 and w_4 are depend on $f_1, \cdots, f_6$ and $f_1, \cdots, f_5$, respectively. So we set

$$w_3 = w_3^{f_1,\cdots,f_6}(\mathbf{x}) \text{ and } w_4 = w_4^{f_1,\cdots,f_5}(\mathbf{x}) ,$$

where $\mathbf{x} = (x_1, x_2, x_3, x_4) \in I_{4n}$ is an input value of KASUMI. As the similar work to previous section, we define bad event needed to prove Theorem 3.

Definition 12. *For every n-bit permutations $f_1, \cdots, f_6$, $BAD(f_1, \cdots, f_6)$ is defined as the set of all D-transcripts $\sigma = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), \cdots, (\mathbf{x}^{(q)}, \mathbf{y}^{(q)})\}$ satisfying: $\exists 1 \le i < j \le q$ such that*

$$w_3^{f_1,\cdots,f_6}(\mathbf{x}^{(i)}) = w_3^{f_1,\cdots,f_6}(\mathbf{x}^{(j)}) \text{ or } w_4^{f_1,\cdots,f_5}(\mathbf{x}^{(i)}) = w_4^{f_1,\cdots,f_5}(\mathbf{x}^{(j)}) .$$

Lemma 5. *Let $f_1, \cdots, f_6$ be chosen independently from UPE $\mathcal{P}_n$. Then for any $\mathcal{D}$-transcript $\sigma = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), \cdots, (\mathbf{x}^{(q)}, \mathbf{y}^{(q)})\}$,*

$$Pr(\sigma \in BAD(f_1, \cdots, f_6)) \le \frac{q^2 - q}{2^n - 1} .$$

Proof. Let $\alpha_k^{(i)}$ be the n-bit input value of $f_k (k = 1, \cdots, 6)$ when the query of $\mathcal{D}$ is $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)}) (i = 1, \cdots, q)$. For example,

$$\begin{aligned} \alpha_3^{(i)} &= x_3^{(i)} \oplus f_1(x_4^{(i)}) , \\ \alpha_5^{(i)} &= x_1^{(i)} \oplus x_3^{(i)} \oplus f_1(x_4^{(i)}) \oplus f_2(x_3^{(i)}) \oplus f_3(x_3^{(i)} \oplus f_1(x_4^{(i)})) . \end{aligned}$$

Then it is easy to show that if $\alpha_k^{(i)} \neq \alpha_k^{(j)}$ for some $k = 1, \cdots, 6$, by Lemma 1

$$Pr\left(w_3^{f_1,\cdots,f_6}(\mathbf{x}^{(i)}) = w_3^{f_1,\cdots,f_6}(\mathbf{x}^{(j)})\right) \le \frac{1}{2^n - 1} ,$$

otherwise ($\alpha_k^{(i)} = \alpha_k^{(j)}$, for all $k = 1, \cdots, 6$) we obtain that this probability is zero, since in this case $w_3^{f_1,\cdots,f_6}(\mathbf{x}^{(i)}) = w_3^{f_1,\cdots,f_6}(\mathbf{x}^{(j)})$ implies to $\mathbf{x}^{(i)} = \mathbf{x}^{(j)}$ which contradicts to the assumption that all queries are distinct. By the similar argument we can also show that $Pr(w_4^{f_1,\cdots,f_5}(\mathbf{x}^{(i)}) = w_4^{f_1,\cdots,f_5}(\mathbf{x}^{(j)})$ has the same upper bound. □

Lemma 6. *Let Λ_{4n} be the $4n$-bit permutation ensemble obtained from the four round KASUMI of Figure 1 where all f_i's$(i = 1, \cdots, 12)$ are independently chosen from the n-bit UPE. Then for any $\mathcal{D}$-transcript $\sigma = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), \cdots, (\mathbf{x}^{(q)}, \mathbf{y}^{(q)})\}$,*

$$|Pr\left(T_{\Lambda_{4n}} = \sigma \mid \ \sigma \notin BAD(f_1, \cdots, f_6)\right) - Pr\left(T_{\mathcal{P}_{4n}} = \sigma\right)| < \varepsilon'_{n,q} \ ,$$

where

$$\varepsilon'_{n,q} = \frac{1}{2^{4n}(2^n - 1)^4 \cdots (2^n - q + 1)^4} .$$

Proof. For any possible $\mathcal{D}$-transcript we have that

$$Pr\left(T_{\mathcal{P}_{4n}} = \sigma\right) = \frac{(2^{4n} - q)!}{2^{4n}!} \ .$$

In Figure 1, by considering four paths $w_3 \rightarrow f_8 \rightarrow z_2$, $w_3 \rightarrow f_8 \rightarrow f_{10} \rightarrow f_{12} \rightarrow z_3$, $w_4 \rightarrow f_7 \rightarrow f_9 \rightarrow z_1$, and $w_4 \rightarrow f_7 \rightarrow f_9 \rightarrow f_{11} \rightarrow z_4$, we can obtain that

$$Pr\left(T_{\Lambda_{4n}} = \sigma \mid \ \sigma \notin BAD(f_1, \cdots, f_6)\right) = \left(\frac{(2^n - q)!}{2^n!}\right)^4 \ ,$$

which complete the proof of this lemma. □

Proof of Theorem 3: From Lemma 5 and 6, Theorem 3 is proved straightforwardly by the similar process in the proof of Theorem 1. □

3 Provable Security of the Encryption Mode $f8$

To guarantee the message confidentiality over the wireless link of W-CDMA for 3GPP, $f8$ encryption mode has been proposed, which is based on the block cipher KASUMI[12]. In this section we examine the provable security of the 3GPP encryption mode $f8$ under the assumption that the underlying block cipher is a pseudorandom permutation. Note that this assumption is reasonable from the result of previous section.

3.1 Notions of Security for a Symmetric Encryption Mode

Symmetric encryption scheme is defined as a triple of algorithms, $(\mathcal{K}, \mathcal{E}, \mathcal{D})$, where $\mathcal{K}$ is the probabilistic algorithm for key generation, $\mathcal{E}$ is the probabilistic algorithm which encrypts the plaintext M with the key K and outputs the ciphertext C, and $\mathcal{D}$ is the the deterministic algorithm which decrypts the ciphertext C with the key K and outputs the corresponding plaintext M. Here M is selected in a set of messages. Bellare et al.[2] considered four notions for security of symmetric encryption modes. "Real-or-random indistinguishability" and its variant "left-or-right indistinguishability" were first introduced. "Find-then-guess security" and "semantic security" which are the notions for the asymmetric encryption scheme, were adapted to the symmetric setting. They also investigated the relation among these notions of security[2]. Real-or-random and left-or-right indistinguishability were equivalent up to a small constant factor in the reduction. Also these notions had a security-preserving reduction to find-then-guess security. However the reduction from find-then-guess security to left-or-right indistinguishability was not security-preserving. It had security-preserving reductions between find-then-guess and semantic security.

Here we analyze the security of 3GPP $f8$ mode by applying the notion of left-or-right indistinguishability, since the left-or-right security implies good reductions to the other three definitions as described above. Left-or-right indistinguishability is a strong form of chosen-plaintext security. It considers two

different games. In either game a query is a pair (x_1, x_2) of equal-length strings from the given message space. In either game a random key $a \in K$ is selected at random and fixed for duration of the game. In Game 1, an oracle receiving (x_1, x_2) responds with $\mathcal{E}_a(x_1)$. In Game 2, it responds with $\mathcal{E}_a(x_2)$. Thus Game 1 provides a "left" oracle and Game 2 provides a "right" oracle. An encryption scheme is secure if a reasonable adversary cannot obtain significant advantage in distinguishing Game 1 and 2.

Definition 13 (Left-or-right indistinguishability[2]). *Encryption scheme $(\mathcal{K}, \mathcal{E}, \mathcal{D})$ is said to be $(t, q, \mu; \epsilon)$-secure, in left-or-right sense, if for any adversary $\mathcal{A}$ who runs in time at most t and makes at most q oracle queries, totaling at most μ bits,*

$$ADV_{\mathcal{A}}^{lr} \stackrel{\text{def}}{=} \left| Pr_{a \leftarrow K}\left(\mathcal{A}^{\mathcal{E}_a(\mathcal{O}(1,(\cdot,\cdot)))} = 1\right) - Pr_{a \leftarrow K}\left(\mathcal{A}^{\mathcal{E}_a(\mathcal{O}(2,(\cdot,\cdot)))} = 1\right) \right| \leq \epsilon \ .$$

Encryption scheme $(\mathcal{K}, \mathcal{E}, \mathcal{D})$ is $(t, q, \mu; \epsilon)$-break, in left-or-right sense, if for an adversary $\mathcal{A}$ who runs in time at most t and makes at most q oracle queries, totaling at most μ bits, $ADV_{\mathcal{A}}^{lr} > \epsilon$.

In the above definition $\mathcal{A}^{\mathcal{E}_a(\mathcal{O}(1,(\cdot,\cdot)))}$ and $\mathcal{A}^{\mathcal{E}_a(\mathcal{O}(2,(\cdot,\cdot)))}$ indicate $\mathcal{A}$ with an oracle $\mathcal{O}$ which returns $y = \mathcal{E}_a(x_1)$ and $y = \mathcal{E}_a(x_2)$, respectively, in response to query (x_1, x_2). And $Pr_{a \leftarrow K}(\mathcal{A}^{\mathcal{E}_a(\mathcal{O}(i,(\cdot,\cdot)))} = 1)$ $(i = 1, 2)$ denotes the probability that the adversary $\mathcal{A}$ with an oracle $\mathcal{O}(i, (\cdot, \cdot))$ $(i = 1, 2)$ outputs 1 when a key a is chosen randomly from the key space K.

The encryption mode $f8$ is based on the block cipher KASUMI and this is a pseudorandom permutation ensemble by referring to last section. Let $\mathcal{B}_l$ be the function family obtained from a block cipher with l-bit input/output values. To analyze the provable security of $f8$ mode, we need more rigorous definition about PPE than Definition 4.

Definition 14. *A permutation family $\mathcal{B}_l$ is said to be a $(t, q; \epsilon)$-secure PPE if for any distinguisher $\mathcal{D}$ who makes at most q oracle queries and runs in time at most t, $ADV_{\mathcal{D}} \leq \epsilon$.*

3.2 Security of $f8$ Encryption Mode

In this subsection, we prove the security of 3GPP $f8$ encryption mode by using the notion of left-or-right security. The underlying function of the encryption mode is fixed to a PPE $\mathcal{B}_l$ with l-bit input/output length. Let $a \in K$ be the key shared between the two parties who run the encryption scheme. It will be used to specify the function $g = \mathcal{B}_l[a]$ and $g' = \mathcal{B}_l[a \oplus KM]$ determined by the key a and $a \oplus KM$, respectively, where KM is an 128-bit fixed constant. We describe rigorously the encryption mode $f8$ as the following scheme. This scheme is also illustrated in Figure 2.

The scheme $f8^g(x)$ works as follows:

Function $f8^g(x)$
  $IV \leftarrow g'(Count||Direction||Bearer||0 \ldots 0)$

$Reg_1 = IV$
for $i = 1, \ldots, n$ do
$o_i = g(Reg_i)$
$y_i = o_i \oplus x_i$
$Reg_{i+1} = IV \oplus i \oplus o_i$
return $(y_1 \ldots y_n)$

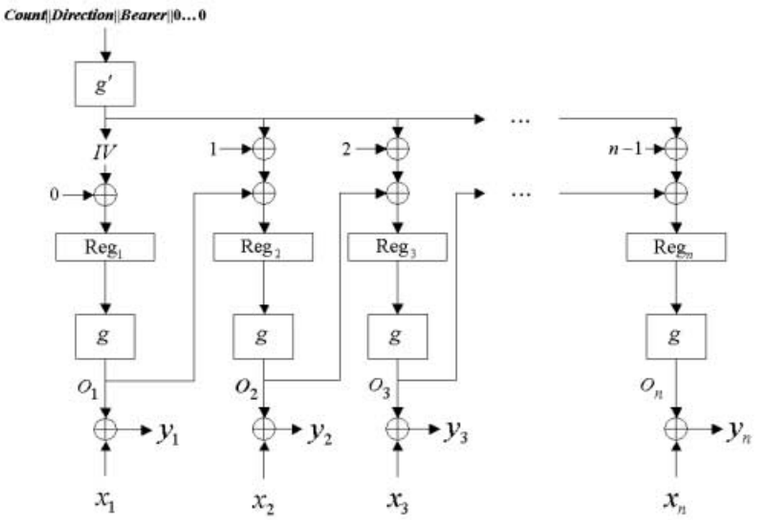

Fig. 2. 3GPP $f8$ encryption mode

In the above scheme *Count* is an encryption sequence number of 32-bit length depending on the time, *Bearer* is a 5-bit bearer identifier, *Direction* is an 1-bit direction identifier, and $0 \ldots 0$ denotes the padding so that the length of the input is an l-bit. The difference between OFB and $f8$ mode is that an initial nonce $ctr = (Count||Direction||Bearer||0 \ldots 0)$ is not sent to the receiver and $g'(ctr)$ is applied to the underlying function g, instead of ctr in a cleartext.

We consider two function family. $f8^{\mathcal{P}_l}$ is the set of all functions $f8^g$, where g is chosen from the UPE $\mathcal{P}_l$, and $f8^{\mathcal{B}_l}$ is the set of all functions $f8^g$, where g is chosen from the PPE $\mathcal{B}_l$. We first derive an upper bound on the success of any adversary trying to break the $f8^{\mathcal{P}_l}$ in the left-or-right sense. Next we examine the security of $f8^{\mathcal{B}_l}$. The basic idea for proving the security of $f8$ is that left-or-right security breaks down at the first repetition of the value of Reg. If $Reg_i = Reg_j$ for $i \neq j$, then also $o_i = o_j$. Hence $y_i \oplus y_j = x_i^b \oplus x_j^b$ $(b = 1, 2)$. Thus b is revealed if $x_i^1 \oplus x_j^1 \neq x_i^2 \oplus x_j^2$.

Lemma 7. *Let $\mathcal{A}$ be any adversary attacking $f8^{\mathcal{P}_l}$ in the left-or-right sense, making at most q queries, totaling at most μ bits. Then*

$$ADV^{lr}_{\mathcal{A}} \leq \delta_{f8^{\mathcal{P}_l}} \stackrel{\text{def}}{=} \frac{\mu/l \cdot (\mu/l - 1)}{2^{l+1}} .$$

Proof. Let $(x^1_1, x^2_1), \ldots, (x^1_q, x^2_q)$ be the oracle queries of the adversary $\mathcal{A}$, each consisting of a pair of equal length messages. These queries are random variables that depend on the coin tosses of $\mathcal{A}$ and responses of the oracle to previous queries. Let $ctr_i = (Count_i||Direction_i||Bearer_i||0\ldots0)$ and $IV_i = g'(ctr_i) \in \{0,1\}^l$, associated to (x^1_i, x^2_i) as computed by the oracle, for $i = 1, \ldots, q$. Let n_i be the number of blocks in the i-th query, $x^b_i = x^b_i[1] \cdots x^b_i[n_i]$ ($b \in \{1,2\}$) be the i-th query message, and $y_i = y_i[1] \cdots y_i[n_i]$ be the response of the oracle to the i-th query message. $Reg_i = Reg_i[1] \cdots Reg_i[n_i]$ is the contents of the register Reg in the i-th query, where $Reg_i[j]$ ($j \in \{1, \ldots, n_i\}$) denotes the content of the register corresponding to the j-th block of the i-th query message. We set $o_i[j]$ is a value computed by applying $Reg_i[j]$ to the function g. Let $Pr_1(\cdot)$ denote the probability in Game 1 providing the adversary $\mathcal{A}$ with the left oracle, and $Pr_2(\cdot)$ denote the probability in Game 2 providing the adversary $\mathcal{A}$ with the right oracle.

Let C be the collision event, i.e., $Reg_i[k] = Reg_j[k']$ whenever $(i,k) \neq (j,k')$, for all $i, j = 1, \ldots, q$ and $k = 1, \ldots, n_i$ and $k' = 1, \ldots, n_j$. The event C^c, complement of C, depends on IV_i, $o_i[k]$ and k for each query. Since g and g' are chosen from the UPE $\mathcal{P}_l$, IV_i and $o_i[k]$ are random and independent of the message given to the oracle. Thus the collision probability does not depend on b, and the following equation holds:

$$Pr_1(C^c) = Pr_2(C^c) . \tag{4}$$

For the same reason, if no collision occurs, the adversary outputs 1 with the same probability for Game 1 and Game 2 because each ciphertext block given to the adversary is independent of any previous ciphertext blocks and of message blocks. Namely, the following holds:

$$Pr_1\left(\mathcal{A} = 1 \mid \ C^c\right) = Pr_2\left(\mathcal{A} = 1 \mid \ C^c\right) . \tag{5}$$

Therefore, by using the equation (4) and (5), we can write the adversary's advantage as follows:

$$\begin{aligned} ADV^{lr}_{\mathcal{A}} &= \left|Pr_1(\mathcal{A} = 1) - Pr_2(\mathcal{A} = 1)\right| \\ &= \big|Pr_1(\mathcal{A} = 1 \mid C) \cdot Pr_1(C) + Pr_1(\mathcal{A} = 1 \mid C^c) \cdot Pr_1(C^c) \\ &\quad - Pr_2(\mathcal{A} = 1 \mid C) \cdot Pr_2(C) - Pr_2(\mathcal{A} = 1 \mid C^c) \cdot Pr_2(C^c)\big| \\ &= \left|(Pr_1(\mathcal{A} = 1 \mid C) - Pr_2(\mathcal{A} = 1 \mid C))Pr_1(C)\right| \\ &\leq Pr_1(C) . \end{aligned}$$

Given the equation (4) we drop the subscript in talking about the probability of C and write the above just as $Pr(C)$. Now we want to compute the upper

bound of $Pr(C)$. The adversary does not know the contents of the register Reg because she does not know $IV_i = g(ctr_i)$. Hence the adversary does not identify the collision $Reg_i[k] = Reg_j[k']$ ($(i,k) \neq (j,k')$ for all $i,j = 1,\ldots,q$ and $k = 1,\ldots,n_i$ and $k' = 1,\ldots,n_j$). However the adversary knows the values $o_i[k]$ since she knows the queried message block $x_i[k]$ and the answered ciphertext block $y_i[k]$. Then she can identity $o_i[k] = o_j[k']$. Since g is a permutation, the output collision, $o_i[k] = o_j[k']$, implies the following:

$$o_i[k] = o_j[k'] \;\Leftrightarrow\; g(Reg_i[k]) = g(Reg_j[k']) \;\Leftrightarrow\; Reg_i[k] = Reg_j[k'] \,.$$

Thus, to compute the upper bound of $Pr(C)$ we compute the probability of the output collision event, T, i.e., $o_i[k] = o_j[k']$ whenever $(i,k) \neq (j,k')$, for all $i,j = 1,\ldots,q$ and $k = 1,\ldots,n_i$ and $k' = 1,\ldots,n_j$. We define the stream B as

$$B = o_1[1]\ldots o_1[n_1]o_2[1]\ldots o_2[n_2]\ldots o_q[1]\ldots o_q[n_q].$$

That is, B is the output values of g until the n_q-th encryption of the last q-th query. The length of B is $Q = l \cdot \sum_{i=1}^{q} n_i \leq \mu$ bits. We first compute the number of streams with a collision $o_i[k] = o_j[k']$ for every possible pair (i,k) and (j,k') $((i,k) \neq (j,k'), 1 \leq i,j \leq q, k = 1,\ldots,n_i, k' = 1,\ldots,n_j)$. As $o_i[k] = o_j[k']$, there are 2^l possible values for the both values. The remaining $Q - 2l$ bits have 2^{Q-2l} possibilities. Thus the number of streams with a collision is 2^{Q-l}. There are $(\mu/l)(\mu/l-1)/2$ possible pairs (i,k) and (j,k'). Hence the number η of streams B with at least one collision is less than $(\mu/l)(\mu/l-1)2^{Q-l-1}$. The stream B has 2^Q possibilities. Thus

$$Pr(T^c) = \frac{(2^Q - \eta)}{2^Q} \geq 1 - \frac{(\mu/l)(\mu/l-1)}{2^{l+1}} \,.$$

This implies the following because of $Pr(C) = Pr(T)$:

$$Pr(C) \leq \frac{(\mu/l)(\mu/l-1)}{2^{l+1}} \,. \quad \square$$

In the practical situation, because the underlying block cipher g is modeled as a pseudorandom permutation, we prove the security of 3GPP $f8$ mode using a pseudorandom permutation. This is derived from the Lemma 7.

Theorem 4. *Let $\mathcal{B}_l$ be a $(t', q'; \epsilon')$-secure PPE with l-bit input/output length. Then $f8^{\mathcal{B}_l}$ scheme is $(t, q, \mu; \epsilon)$-secure in the left-or-right sense. Here $q = q'$, $\mu = q'l$, $t = t' - c\frac{\mu}{l}(l+l)$ and $\epsilon = 2\epsilon' + \delta_{f8^{\mathcal{P}_l}}$, where $c > 0$ is a small constant and $\delta_{f8^{\mathcal{P}_l}} = \frac{(\mu/l)(\mu/l-1)}{2^{l+1}}$.*

Proof. The details of this proof are omitted since it is similar to the proof of Theorem 12 in [2] by replacing pseudorandom function with pseudorandom permutation. $\square$

4 Conclusion

In this work we examined the pseudorandomness of the block cipher KASUMI and the provable security of $f8$. We proved that FI function within KASUMI composed of four round unbalanced MISTY-type structure was a pseudorandom permutation. And we showed that the three round KASUMI was not a permutation ensemble but the four round KASUMI was a pseudorandom permutation ensemble under the adaptive distinguisher model. Moreover we provided the upper bound on the security of $f8$ encryption mode under the reasonable assumption from the first result by means of the left-or-right security notion.

References

1. M. Bellare, J. Kilian, and P. R. Rogaway, *The security of cipher block chaining message authentication codes*, Advances in Cryptology-Crypto '94, LNCS 839, Springer-Verlag, 1994, pp. 341-358.
2. M. Bellare, A. Desai, E. Jokipii, and P. Rogaway, *A Concrete Security Treatment of Symmetric Encryption: Analysis of the DES Modes of Operation*, 38th Symposium on Foundations of Computer Science(FOCS), IEEE Computer Society, 1997, pp. 394-403.
3. FIPS PUB 81, DES Modes of Operation, Federal Information Processing Standards Publication 81, December 2, 1980.
4. H. Gilbert and M. Minier, *New results on the pseudorandomness of some block cipher constructions*, Preproceedings of Fast Software Encryption workshop 2001, (2001, Yokohama), pp. 260-277.
5. T. Iwata, T. Yoshino, T. Yuasa, and K. Kurosawa, *Round security and super-pseudorandomness of MISTY type structure*, Preproceedings of Fast Software Encryption workshop 2001, (2001, Yokohama), pp. 245-259.
6. J. S. Kang, O. Y. Yi, D. W. Hong, and H. S. Cho, *Pseudorandomness of MISTY-type transformations and the block cipher KASUMI*, ACISP2001, LNCS 2119, Springer-Verlag, 2001, pp. 60-73.
7. M. Luby and C. Rackoff, *How to construct pseudorandom permutations and pseudorandom functions*, SIAM J. Comput., Vol. 17, 1988, pp. 189-203.
8. M. Matsui, *New permutation of Block Ciphers with Provable Security against Differential and Linear Cryptalaysis*, Fast Software Encryption, LNCS 1039, Springer-Verlag, 1996, pp. 205-218.
9. M. Matsui, *New Block Encryption Algorithm MISTY*, Fast Software Encryption'97, LNCS 1267, Springer-Verlag, 1997, pp. 54-68.
10. M. Naor and O. Reingold, *On the construction of pseurandom permutations: Luby-Rackoff revisited*, J. Cryptology, Vol. 12, 1999, pp. 29-66.
11. K. Sakurai and Y. Zheng, *On non-pseudorandomness from block ciphers with provable immunity against linear cryptanaysis*, IEICE Trans. Fundamentals, Vol. E80-A, No. 1, 1997, pp. 19-24.
12. 3G TS 35.201, *Specification of the 3GPP confidentiality and integrity algorithm; Document 1: f8 and f9 specifications*, available at http://www.3gpp.org
13. A. Alkassar, A. Geralay, B. Pfitzmann, and A. R. Sadeghi, *Optimized Self-Synchronizing Mode of Operation*, Preproceedings of 8th Fast Software Encryption Workshop, April 2, 2001, pp. 82-96.