Security of the MISTY Structure in the Luby-Rackoff Model: Improved Results

Gilles Piret and Jean-Jacques Quisquater

UCL Crypto Group
Place du Levant, 3, B-1348 Louvain-la-Neuve, Belgium
{piret, jjq}@dice.ucl.ac.be

Abstract. In this paper we consider the security of the Misty structure in the Luby-Rackoff model, if the inner functions are replaced by involutions without fixed point. In this context we show that the success probability in distinguishing a 4-round L-scheme from a random function is $O(m^2/2^n)$ (where m is the number of queries and $2n$ the block size) when the adversary is allowed to make adaptively chosen encryption queries. We give a similar bound in the case of the 3-round R-scheme. Finally, we show that the advantage in distinguishing a 5-round scheme from a random permutation when the adversary is allowed to adaptively chosen encryption as well as decryption queries is also $O(m^2/2^n)$. This is to our knowledge the first time involutions are considered in the context of the Luby-Rackoff model.

1 Introduction.

Proving the security of block ciphers has been a long-standing problem, and it is not solved yet. In their seminal paper [4], M. Luby and C. Rackoff introduced a model that permits the assessment of the security of some block cipher constructions. In this model, only the high-level structure of a block cipher is considered, while the lower-level operations are replaced by random functions. This last hypothesis is pretty strong, but at least it permits to guarantee that the basic structure of a block cipher is not flawed from the beginning.

More precisely, the model works as follows: let $\Phi(f_1, ..., f_r)$ be a construction which to r functions $f_1, ..., f_r : \{0,1\}^n \rightarrow \{0,1\}^n$ associates one function $F : \{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$. We consider a distinguishing algorithm $\mathcal{A}$ which has unbounded computation capabilities, and can make a certain number of adaptively chosen encryption queries to an oracle function $\mathcal{O} : \{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$ he received as an input[1]. Based on the answers he

[1] The size of the input and output spaces of $\mathcal{O}$ are often $2n$ bits, where n is the size of the inner functions. However these constraints are absolutely not mandatory; the input and output sizes do not even need to be the same.

H. Handschuh and A. Hasan (Eds.): SAC 2004, LNCS 3357, pp. 100–115, 2005.

obtains to his queries, $\mathcal{A}$ outputs either 0 or 1. Let $p = \Pr[\mathcal{A}^{\Phi(f_1^*,...,f_r^*)} = 1]$ and $p^* = \Pr[\mathcal{A}^{F^*} = 1]$ denote the probability that $\mathcal{A}$ outputs 1 when $\mathcal{O}$ is respectively a function of the form $\Phi(f_1^*, ..., f_r^*)$, where $f_1^*, ..., f_r^*$ are ***perfect*** random functions (i.e. functions randomly chosen with respect to the uniform distribution), or $\mathcal{O}$ itself is a perfect random function F^*. We are interested in the *advantage* $\mathcal{A}$ has in distinguishing $\Phi(f_1^*, ..., f_r^*)$ from F^*: $\mathrm{Adv}_{\mathcal{A}}(\Phi(f_1^*, ..., f_r^*), F^*) = |p - p^*|$. A security proof in the Luby-Rackoff model consists in upper bounding this advantage (as a function of the number of queries m and the block size $2n$) for all possible distinguishers $\mathcal{A}$. If for n big enough, and for all distinguishing algorithms $\mathcal{A}$ of which the number of queries m is polynomial in n, $\mathrm{Adv}_{\mathcal{A}}$ is polynomially small, then Φ is said to be ***pseudorandom***. If this criteria still holds when *decryption* queries are allowed as well, then Φ is said to be ***superpseudorandom***. As a shortcut, an algorithm allowed to make adaptative encryption queries only will often be called ***pseudorandom distinguisher***, and an algorithm allowed to make both adaptative encryption and adaptative decryption queries will be called ***superpseudorandom distinguisher***.

Luby and Rackoff's paper initiated a significant amount of research in the area: in 1992 Patarin [11, 12] made explicit the link between the advantage and the transition probability associated with a given structure Φ (see section 2.3); this gives a practical way of upper bounding the advantage. The same year, Maurer showed how to generalise undistinguishability results to *locally* random functions. More recently, Ramzan and Reyzin introduced a new model which assumes that the attacker has oracle access to some of the round functions [16]. Besides, the Feistel structure (first examined by Luby and Rackoff) was widely studied. On the one hand, its security bounds were tried to be improved [11, 13, 14, 15]. On the other hand, slightly modified constructions were examined: constructions were some of the round functions are identical [12], or are replaced by hash functions for example [5, 10]. Moreover some other constructions were also examined [9, 19].

Recently, constructions used in the block ciphers Misty [6] and Kasumi were examined. In 1997, Sakurai and Zheng [17] presented several negative results (i.e. non-pseudorandomness and non-superpseudorandomness) on these schemes. Then Gilbert and Minier [8] showed in 2001 that the 4-round Misty construction (called *L-scheme*) is pseudorandom, while 3 rounds of its inverse (called *R-scheme*) is sufficient to obtain pseudorandomness. Moreover they showed that 5 rounds of these constructions are necessary to obtain superpseudorandomness. The same year, Iwata et al. [3] showed that some of the 5 inner permutations can be replaced

by uniform ϵ-XOR universal permutations without losing superpseudorandomness; moreover, following the model of Ramzan and Reyzin [16], they show that oracle access to some specific inner permutations does not change superpseudorandomness either. Finally, the next year about the same authors showed that the second inner permutation of a 5-round Misty does not need to be cryptographic at all to guarantee superpseudorandomness: it can be a constant and public transformation g, provided g satisfies $g(x) \oplus x \neq g(x') \oplus x'$ [2].

In this paper, we consider another restriction on the inner functions: namely, we assume that all of them are random involutions (i.e. permutations c such that $\forall x : c(c(x)) = x$) without fixed point. For implementation reasons, involutions were a basis of the design of several recent block ciphers (see e.g. Khazad [1], Anubis, Noekeon, `ICEBERG` [18]), hence the interest of such hypothesis. We show that the pseudorandom character of Misty constructions is preserved under this constraint (the number of rounds considered remaining unchanged).

2 Preliminaries.

2.1 The Misty L- and R-Schemes.

We describe two basic schemes: the L-scheme has been used in the Misty [6] and Kasumi block ciphers, the R-scheme is almost its inverse (we follow the terminology used by Gilbert and Minier [8]).

We define a 1-round L-scheme as a $2n$-bit permutation ψ_L taking a n-bit permutation c as a round function and such that:

$$\psi_L(c)(L, R) = (R, c(L) \oplus R)$$

It is depicted in Figure 1. An r-round L-scheme is simply the composition of r 1-round L-schemes, transforming r n-bit permutations $c_1, ..., c_r$ into a $2n$-bit permutation:

$$\psi_L(c_1, c_2, ..., c_r) = \psi_L(c_r) \circ ... \circ \psi_L(c_1)$$

A 1-round R-scheme transforms a n-bit permutation c into a $2n$-bit permutation $\psi_R(c)$ too. It is defined as (see Figure 1):

$$\psi_R(c)(L, R) = (c(L) \oplus R, c(L))$$

The composition of r 1-round R-schemes is a r-round R-scheme:

$$\psi_R(c_1, c_2, ..., c_r) = \psi_R(c_r) \circ ... \circ \psi_R(c_1)$$

In this paper we consider variants of the ψ_L and ψ_R schemes, where the last XOR operation is omitted, as well as the last swap. We call them ψ'_L and ψ'_R.

Remark 1. Cryptographically speaking, ψ'_L and ψ'_R are equivalent respectively to ψ_L and ψ_R.

Remark 2. $\psi'_L(c_1, c_2, ..., c_r)$ and $\psi'_R(c_r^{-1}, c_{r-1}^{-1}, ..., c_1^{-1})$ are inverses of each other. It implies that their security against superpseudorandom distinguishers is the same.

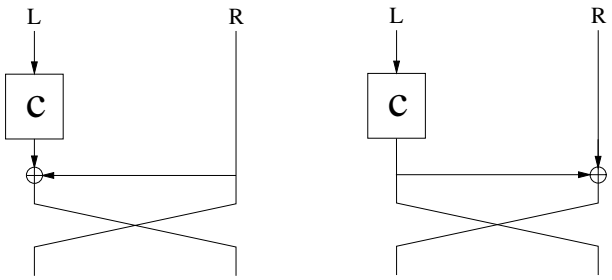

Fig. 1. 1-round L-scheme at left, 1-round R-scheme at right

2.2 Notations.

Throughout this paper we use the following notations:

- I_n denotes the $\{0,1\}^n$ set.
- $I := I_n^m$ (where m is the number of plaintext-ciphertext pairs considered).
- For $\mathbf{X}, \mathbf{Y} \in I$: $\mathbf{X} \sim \mathbf{Y}$ informally means that $\mathbf{X}$ and $\mathbf{Y}$ could be the inputs and outputs of a permutation. More formally: $\forall i, j \in [1...m] : X_i = X_j \Leftrightarrow Y_i = Y_j$.
- $I^{\neq} := \{\mathbf{X} \in I | \nexists i \neq j \in [1...m] : X_i = X_j\}$ $\quad I^{=} := I \backslash I^{\neq}$.
- Let $\mathcal{X}$ be the subset of I_{2n}^m such that $\forall((X_i, Y_i))_{i\in[1..m]} \in \mathcal{X} : \forall i \neq j : (X_i, Y_i) \neq (X_j, Y_j)$. Then the m inputs to ψ_L (or ψ_R) are assumed[2] to belong to $\mathcal{X}$ and denoted by $(\mathbf{L}, \mathbf{R}) = ((L_i)_{i\in[1..m]}, (R_i)_{i\in[1..m]}) \in \mathcal{X}$. Similarly the m corresponding outputs are denoted by $(\mathbf{S}, \mathbf{T}) = ((S_i, T_i))_{i\in[1..m]} \in \mathcal{X}$.

[2] This hypothesis reflects the fact that the distinguisher is assumed not to make two times the same query. As the distinguisher would learn nothing more when repeating a query, there is no loss of generality.

– f^* always denotes a perfect random function (or permutation, or involution without fixed point, depending on the context), i.e. one which is chosen in accordance with the uniform probability distribution.

2.3 Patarin's Coefficient H Technique.

Let $\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})}$ be the probability for a structure $\Phi(f_1,...,f_r)$ to be such that $\Phi(f_1,...,f_r)(\mathbf{L},\mathbf{R}) = (\mathbf{S},\mathbf{T})$ (computed over all possible $f_1,...,f_r$)). Not surprisingly, this probability plays a big role in upper bounding the advantage an algorithm $\mathcal{A}$ has in distinguishing Φ from a perfect random function F^*. The link between $\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})}$ and the best advantage has been quantified by Patarin [11, 12][3]:

Theorem 1 (Patarin). *Let $F : I_{2n} \to I_{2n}$ be a random function; let $F^* : I_{2n} \to I_{2n}$ be a perfect random function. Let m be an integer. If there exists a subset $\mathcal{Y}$ of I_{2n}^m and two positive real numbers ϵ_1 and ϵ_2 such that*

1. $|\mathcal{Y}| > (1-\epsilon_1) \cdot |I_{2n}|^m$
2. $\forall(\mathbf{L},\mathbf{R}) \in \mathcal{X} \quad \forall(\mathbf{S},\mathbf{T}) \in \mathcal{Y} : \mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})} \geq (1-\epsilon_2) \cdot \frac{1}{|I_{2n}|^m}$

Then for any distinguisher $\mathcal{A}$ using m encryption queries

$$Adv_{\mathcal{A}}(F, F^*) \leq \epsilon_1 + \epsilon_2$$

Theorem 1 deals with pseudorandom distinguishers. A similar theorem holds for superpseudorandom distinguishers:

Theorem 2 (Patarin). *Let $C : I_{2n} \to I_{2n}$ be a random permutation; let $C^* : I_{2n} \to I_{2n}$ be a perfect random permutation. Let m be an integer, and $\epsilon > 0$. If for all $(\mathbf{L},\mathbf{R}) \in \mathcal{X}$, and all $(\mathbf{S},\mathbf{T}) \in \mathcal{X}$: $\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})} \geq (1-\epsilon) \cdot \frac{1}{|I_{2n}|^m}$ then for any distinguisher $\mathcal{A}$ using m encryption or decryption queries: $Adv_{\mathcal{A}}(C, C^*) \leq \epsilon + \frac{m(m-1)}{2 \cdot 2^{2n}}$*

3 The 4-Round L-Scheme.

We consider a 4-round L-scheme were the inner permutations $c_1^*, ..., c_4^*$ are perfect random involutions without fixed point. In section 4 we will prove the following lemma:

[3] We particularized Patarin's theorem to the case where the input and output sizes are both $2n$, but it holds for any size.

Lemma 1. *Let $m, n > 0$. Let $(\mathbf{L}, \mathbf{R}) \in \mathcal{X} \subset I_{2n}^m, (\mathbf{S}, \mathbf{T}) \in I \times I^{\neq}$. Then the probability for a 4-uple (c_1, c_2, c_3, c_4) of involutions without fixed point to satisfy $\psi'_L(c_1, ..., c_4)(\mathbf{L}, \mathbf{R}) = (\mathbf{S}, \mathbf{T})$ is lower bounded by*

$$\left[1 - \frac{15m^2}{2^n} - \frac{9}{32}\sum_{k=2}^{\infty}\left(\frac{16m^2}{2^n}\right)^k\right] \cdot \frac{1}{2^{2nm}} \geq \left(1 - \frac{24m^2}{2^n}\right) \cdot \frac{1}{2^{2nm}}$$

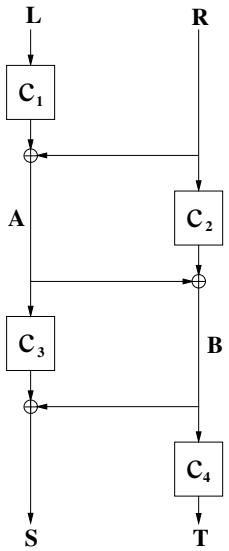

Fig. 2. 4 rounds L-scheme

It allows to prove the following theorem:

Theorem 3. *Let $c_1^*, ..., c_4^*$ be independent perfect random involutions without fixed point on I_n. Let $C := \psi_L(c_1^*, ..., c_4^*)$. Let $F^* : I_{2n} \to I_{2n}$ be a perfect random function. Then for any pseudorandom distinguisher $\mathcal{A}$ allowed to make m queries, we have:*

$$Adv_{\mathcal{A}}(C, F^*) \leq \frac{31m^2}{2 \cdot 2^n} + \frac{9}{32}\sum_{k=2}^{\infty}\left(\frac{16m^2}{2^n}\right)^k \leq \frac{49m^2}{2 \cdot 2^n}$$

Thus $\psi_L(c_1^*, ..., c_4^*)$ is pseudorandom, and secure as long as $m \ll 2^{n/2}$.

Proof. It is an immediate application of theorem 1. The constraint $\mathbf{T} \in I^{\neq}$ in lemma 1 implies a non-zero ϵ_1. More precisely, ϵ_1 is equal to the

probability for a (perfect) random $\mathbf{T} \in I$ to belong to $I^{=}$. It can be shown to be smaller than $\frac{m^2}{2 \cdot 2^n}$:

$$\Pr[\mathbf{T} \in I^{=}] = \Pr[\bigvee_{i<j} T_i = T_j] \leq \sum_{i<j} \Pr[T_i = T_j] \leq \frac{m^2}{2 \cdot 2^n}$$

Lemma 1 gives the corresponding ϵ_2.

4 Proof of Lemma 1.

For a given $(\mathbf{L}, \mathbf{R}, \mathbf{S}, \mathbf{T})$, we define λ and ρ as the number of independent equalities of the form $L_i = L_j$ and $R_i = R_j$ $(i \neq j)$, respectively. We also define two intermediate states during the computation of $\psi'_L(c_1, ..., c_4)$, namely $\mathbf{A} := c_1(\mathbf{L}) \oplus \mathbf{R}$ and $\mathbf{B} := c_2(\mathbf{R}) \oplus \mathbf{A}$ (see Figure 2). Let $\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})}$ be the probability that a random 4-uple (c_1, c_2, c_3, c_4) is such that $\psi'_L(c_1, c_2, c_3, c_4)(\mathbf{L}, \mathbf{R}) = (\mathbf{S}, \mathbf{T})$. Then

$$\begin{aligned} \mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})} = \sum_{\mathbf{A},\mathbf{B} \in I} & \Pr[(c_1(\mathbf{L}) \oplus \mathbf{R} = \mathbf{A}) \wedge (c_2(\mathbf{R}) \oplus \mathbf{A} = \mathbf{B}) \\ & \wedge (c_3(\mathbf{A}) \oplus \mathbf{B} = \mathbf{S}) \wedge (c_4(\mathbf{B}) = \mathbf{T})] \end{aligned} \tag{1}$$

We consider the following conditions $(\mathcal{C})$ on $(\mathbf{A}, \mathbf{B})$:

$(\mathcal{C}1)$ $\mathbf{A} \oplus \mathbf{R} \sim \mathbf{L}$ and $\nexists i, j$ s.t. $L_i = A_j \oplus R_j$.
$(\mathcal{C}2)$ $\mathbf{A} \oplus \mathbf{B} \sim \mathbf{R}$ and $\nexists i, j$ s.t. $R_i = A_j \oplus B_j$.
$(\mathcal{C}3)$ $\mathbf{B} \oplus \mathbf{S} \in I^{\neq}$ and $\nexists i, j$ s.t. $A_i = B_j \oplus S_j$.
$(\mathcal{C}4)$ $\nexists i, j$ s.t. $B_i = T_j$.

Then equation (1) implies:

$$\begin{aligned} \mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})} \geq \sum_{\substack{\mathbf{A},\mathbf{B} \in I^{\neq} \\ (\mathbf{A},\mathbf{B}) \text{ satisfies } (\mathcal{C})}} & \Pr[(c_1(\mathbf{L}) \oplus \mathbf{R} = \mathbf{A})] \cdot \Pr[c_2(\mathbf{R}) \oplus \mathbf{A} = \mathbf{B}] \\ & \cdot \Pr[c_3(\mathbf{A}) \oplus \mathbf{B} = \mathbf{S}] \cdot \Pr[c_4(\mathbf{B}) = \mathbf{T}] \end{aligned} \tag{2}$$

The number of $\mathbf{A}$ such that $(\mathcal{C}1)$ is satisfied is $\frac{(2^n - m + \lambda)!}{(2^n - 2m + 2\lambda)!}$. For a (perfect) random such $\mathbf{A}$ we have:

$$\Pr[\mathbf{A} \in I^{\neq} | (\mathcal{C}1)] \geq 1 - \sum_{i<j} \Pr[A_i = A_j | (\mathcal{C}1)] \tag{3}$$

Consider given $1 \leq i < j \leq m$, and assume $L_i \neq L_j$ and $R_i \neq R_j$. As there are $(2^n-m+\lambda)(2^n-m+\lambda-1)$ possible values for (A_i, A_j) satisfying $(\mathcal{C}_1)$, among which $2^n - m + \lambda$ satisfy $A_i = A_j$, we get

$$\Pr[A_i = A_j|(\mathcal{C}_1)] = \frac{2^n - m + \lambda}{(2^n - m + \lambda)(2^n - m + \lambda - 1)} \leq \frac{2}{2^n} \tag{4}$$

If $L_i = L_j$ or $R_i = R_j$, it is easy to see that $\Pr[A_i = A_j|(\mathcal{C}1)] = 0$.

Then we have

$$\Pr[\mathbf{A} \in I^{\neq}|(\mathcal{C}1)] \geq 1 - \frac{m(m-1)}{2} \cdot \frac{2}{2^n} \geq 1 - \frac{m^2}{2^n} \tag{5}$$

Similarly, the number of $\mathbf{B}$ such that $(\mathcal{C}2)$ is satisfied is $\frac{(2^n-m+\rho)!}{(2^n-2m+2\rho)!}$, and $\Pr[\mathbf{B} \in I^{\neq}|(\mathcal{C}2)] \geq 1 - \frac{m^2}{2^n}$. Finally for a (perfect) random $(\mathbf{A}, \mathbf{B})$ we compute:

$$\begin{aligned}
&\Pr[\mathbf{B} \text{ satisfies } (\mathcal{C}3) \wedge \mathbf{B} \text{ satisfies } (\mathcal{C}4) \wedge \mathbf{A} \in I^{\neq} \wedge \mathbf{B} \in I^{\neq}|(\mathcal{C}1) \wedge (\mathcal{C}2)] \\
&\geq 1 - \Pr[\bigvee_{i<j} B_i \oplus S_i = B_j \oplus S_j|(\mathcal{C}1) \wedge (\mathcal{C}2)] \\
&- \Pr[\bigvee_{i,j} A_i = B_j \oplus S_j|(\mathcal{C}1) \wedge (\mathcal{C}2)] - \Pr[\bigvee_{i,j} B_i = T_j|(\mathcal{C}1) \wedge (\mathcal{C}2)] - 2 \cdot \frac{m^2}{2^n} \\
&\geq 1 - \frac{m(m-1)}{2} \cdot \frac{2}{2^n} - \frac{4m^2}{2^n} - 2 \cdot \frac{m^2}{2^n} \geq 1 - \frac{7m^2}{2^n}
\end{aligned}$$

Thus the number of $(\mathbf{A}, \mathbf{B}) \in I^{\neq}$ satisfying $(\mathcal{C})$ can be lower bounded by:

$$\frac{(2^n - m + \lambda)!}{(2^n - 2m + 2\lambda)!} \cdot \frac{(2^n - m + \rho)!}{(2^n - 2m + 2\rho)!} \cdot (1 - \frac{7m^2}{2^n}) \tag{6}$$

Under these conditions on $(\mathbf{A}, \mathbf{B})$ we can evaluate

$$\Pr[(c_1(\mathbf{L}) \oplus \mathbf{R} = \mathbf{A})] \cdot \Pr[c_2(\mathbf{R}) \oplus \mathbf{A} = \mathbf{B}] \cdot \Pr[c_3(\mathbf{A}) \oplus \mathbf{B} = \mathbf{S}] \cdot \Pr[c_4(\mathbf{B}) = \mathbf{T}]$$

and we obtain:

$$\frac{\frac{(2^n-2m+2\lambda)!}{2^{2^{n-1}-m+\lambda} \cdot (2^{n-1}-m+\lambda)!} \cdot \frac{(2^n-2m+2\rho)!}{2^{2^{n-1}-m+\rho} \cdot (2^{n-1}-m+\rho)!} \cdot \left[\frac{(2^n-2m)!}{2^{2^{n-1}-m} \cdot (2^{n-1}-m)!}\right]^2}{\left[\frac{2^n!}{2^{2^{n-1}} \cdot (2^{n-1})!}\right]^4} \tag{7}$$

After multiplication of (7) by the number of terms (6):

$$\begin{aligned}
&\frac{2^{4m-\lambda-\rho} \cdot (2^n - m + \lambda)! \cdot (2^n - m + \rho)!}{(2^{n-1} - m + \lambda)! \cdot (2^{n-1} - m + \rho)!} \cdot \frac{(2^{n-1})!^4}{(2^n)!^4} \cdot \left[\frac{(2^n - 2m)!}{(2^{n-1} - m)!}\right]^2 \cdot \left(1 - \frac{7m^2}{2^n}\right) \\
&= 2^{4m-\lambda-\rho} \cdot \frac{\prod_{i=0}^{m-\lambda-1} \frac{2^{n-1}-i}{2^n-i} \cdot \prod_{i=0}^{m-\rho-1} \frac{2^{n-1}-i}{2^n-i} \cdot \left(\prod_{i=0}^{m-1} \frac{2^{n-1}-i}{2^n-i}\right)^2}{\left(\prod_{i=m}^{2m-1} 2^n - i\right)^2} \cdot \left(1 - \frac{7m^2}{2^n}\right)
\end{aligned}$$

By lower bounding the products, this expression can be shown to be greater or equal than:

$$2^{4m-\lambda-\rho} \cdot \left(\frac{2^{n-1}-m}{2^n-m}\right)^{4m-\lambda-\rho} \cdot \frac{1}{2^{2nm}} \cdot \left(1-\frac{7m^2}{2^n}\right) \tag{8}$$

It is easy to show that $\frac{2^{n-1}-m}{2^n-m} = \frac{1}{2} - \frac{1}{2}\sum_{k=1}^{\infty}\frac{m^k}{2^{nk}}$. Then (8) is greater or equal than:

$$\left(1-\sum_{k=1}^{\infty}\frac{m^k}{2^{nk}}\right)^{4m} \cdot \frac{1}{2^{2nm}} \cdot \left(1-\frac{7m^2}{2^n}\right) \tag{9}$$

By evaluating the first factor using the binomial theorem, we can show

$$\left(1-\sum_{k=1}^{\infty}\frac{m^k}{2^{nk}}\right)^{4m} \geq 1-\frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{16m^2}{2^n}\right)^k \tag{10}$$

Finally, immediate calculations show that (9) is greater or equal than:

$$\left[1-\frac{15m^2}{2^n}-\frac{9}{32}\sum_{k=2}^{\infty}\left(\frac{16m^2}{2^n}\right)^k\right] \cdot \frac{1}{2^{2nm}} \geq \left(1-\frac{24m^2}{2^n}\right) \cdot \frac{1}{2^{2nm}} \tag{11}$$

which concludes the proof.

5 The 3-Round R-Scheme.

A result similar to theorem 3 can be proved for a 3-round R-scheme:

Theorem 4. *Let c_1^*, c_2^*, c_3^* be independent perfect random involutions without fixed point on I_n. Let $C := \psi_R(c_1^*, c_2^*, c_3^*)$. Let $F^* : I_{2n} \rightarrow I_{2n}$ be a perfect random function. Then for any pseudorandom distinguisher $\mathcal{A}$ allowed to make m queries, we have:*

$$Adv_{\mathcal{A}}(C, F^*) \leq \frac{11m^2}{2^n} + \frac{5}{8}\sum_{k=2}^{\infty}\left(\frac{8m^2}{2^n}\right)^k \leq \frac{13m^2}{2^n}$$

Thus $\psi_R(c_1^*, c_2^*, c_3^*)$ is pseudorandom, and secure as long as $m \ll 2^{n/2}$.

6 The 5-Round Scheme.

The following lemma is proved in the next section:

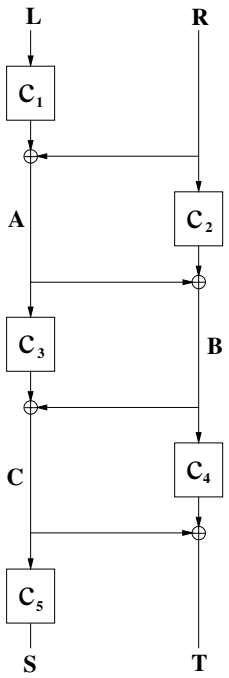

Fig. 3. 5 rounds L-scheme

Lemma 2. *Let $m, n > 0$. Let $(\mathbf{L}, \mathbf{R}), (\mathbf{S}, \mathbf{T}) \in \mathcal{X} \subset I_{2n}^m$. Then the probability for a 5-uple $(c_1, c_2, c_3, c_4, c_5)$ of involutions without fixed point to satisfy $\psi'_L(c_1, ..., c_5)(\mathbf{L}, \mathbf{R}) = (\mathbf{S}, \mathbf{T})$ is lower bounded by*

$$\left(1 - \frac{12m^2}{2^n}\right) \cdot \frac{1}{2nm}$$

Using theorem 2, it implies superpseudorandomness for a 5-round scheme:

Theorem 5. *Let $c_1^*, c_2^*, ..., c_5^*$ be independent perfect random involutions without fixed point of I_n. Let C^* be a perfect random permutation of I_{2n}. Let $C := \psi_L(c_1^*, c_2^*, ..., c_5^*)$ (resp. $C := \psi_R(c_1^*, c_2^*, ..., c_5^*)$). Then for any superpseudorandom distinguisher $\mathcal{A}$ allowed to make m queries:*

$$Adv_{\mathcal{A}}(C, C^*) \leq \frac{12m^2}{2^n} + \frac{m^2}{2 \cdot 2^n}$$

Thus $\psi(c_1^*, c_2^*, ..., c_5^*)$ is superpseudorandom, and secure as long as $m \ll 2^{n/2}$.

The proof of lemma 2 will require the following lemma. Proving it is easy, it is why we do not give the proof here.

Lemma 3. *Let $x, y \in I_n, 0 \neq \Delta \in I_n$. The probability for a random involution without fixed point c to satisfy*

$$c(x) \oplus c(y) = \Delta$$

is at most $4/2^n$.

7 Proof of Lemma 2.

We use the intermediate states $\mathbf{A} := c_1(\mathbf{L}) \oplus \mathbf{R}$, $\mathbf{B} := c_2(\mathbf{R}) \oplus \mathbf{A}$ and $\mathbf{C} := c_3(\mathbf{A}) \oplus \mathbf{B}$ (see Figure 3). Let $\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})}$ be the probability that a random 5-uple $(c_1, c_2, c_3, c_4, c_5)$ of involutions is such that $\psi'_L(c_1, c_2, c_3, c_4, c_5)(\mathbf{L}, \mathbf{R}) = (\mathbf{S}, \mathbf{T})$. Then:

$$\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})} = \sum_{\mathbf{A},\mathbf{B},\mathbf{C} \in I} \Pr[(c_1(\mathbf{L}) \oplus \mathbf{R} = \mathbf{A}) \wedge (c_2(\mathbf{R}) \oplus \mathbf{A} = \mathbf{B}) \wedge (c_3(\mathbf{A}) \oplus \mathbf{B} = \mathbf{C}) \wedge (c_4(\mathbf{B}) \oplus \mathbf{C} = \mathbf{T}) \wedge (c_5(\mathbf{C}) = \mathbf{S})] \tag{12}$$

We define the following three conditions ($\mathcal{C}$) on $(\mathbf{A}, \mathbf{B}, \mathbf{C})$:

($\mathcal{C}1$) $\nexists i, j : L_i = A_j \oplus R_j$ and $\nexists i, j : R_i = A_j \oplus B_j$
($\mathcal{C}2$) $\nexists i, j : A_i = B_j \oplus C_j$ and $\nexists i, j : B_i = C_j \oplus T_j$
($\mathcal{C}3$) $\nexists i, j : C_i = S_j$

Then $\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})}$ is greater or equal than

$$\sum_{\substack{\mathbf{A},\mathbf{B} \in I^{\neq} \\ \mathbf{A},\mathbf{B} \text{ satisfy } (\mathcal{C}1)}} \Big(\Pr[(c_1(\mathbf{L}) \oplus \mathbf{R} = \mathbf{A}) \wedge (c_2(\mathbf{R}) \oplus \mathbf{A} = \mathbf{B})] \cdot \sum_{\substack{\mathbf{C} \in I \\ \mathbf{C} \text{ satisfies } (\mathcal{C}2),(\mathcal{C}3)}} \Pr[c_3(\mathbf{A}) \oplus \mathbf{B} = \mathbf{C}] \cdot \Pr[c_4(\mathbf{B}) \oplus \mathbf{C} = \mathbf{T}] \cdot \Pr[c_5(\mathbf{C}) = \mathbf{S}] \Big) \tag{13}$$

We first evaluate the inner sum for given $\mathbf{A}, \mathbf{B} \in I^{\neq}$ satisfying ($\mathcal{C}1$). Adding constraints $\mathbf{C} \sim \mathbf{S}$, $\mathbf{C} \oplus \mathbf{T} \in I^{\neq}$ and $\mathbf{B} \oplus \mathbf{C} \in I^{\neq}$ only removes zero terms from the sum. Thus it is equal to:

$$\sum_{\substack{\mathbf{C} \sim \mathbf{S} \\ \mathbf{C} \oplus \mathbf{T} \in I^{\neq},\ \mathbf{B} \oplus \mathbf{C} \in I^{\neq} \\ \mathbf{C} \text{ satisfies } (\mathcal{C}2),(\mathcal{C}3)}} \Pr[c_3(\mathbf{A}) \oplus \mathbf{B} = \mathbf{C}] \cdot \Pr[c_4(\mathbf{B}) \oplus \mathbf{C} = \mathbf{T}] \cdot \Pr[c_5(\mathbf{C}) = \mathbf{S}] \tag{14}$$

It is easy to see that $|\{\mathbf{C} \in I : \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)\}| = \frac{(2^n-m+\sigma)!}{(2^n-2m+2\sigma)!}$. Moreover we compute:

$$\begin{aligned}
&\Pr[\mathbf{C} \oplus \mathbf{T} \in I^{\neq} \wedge \mathbf{B} \oplus \mathbf{C} \in I^{\neq} \wedge (\mathcal{C}_2) | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] \\
&\geq 1 - \sum_{i<j} \Pr[C_i \oplus T_i = C_j \oplus T_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] \\
&- \sum_{i<j} \Pr[B_i \oplus C_i = B_j \oplus C_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)]) \\
&- \sum_{i,j} \Pr[A_i = B_j \oplus C_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] - \sum_{i,j} \Pr[B_i = C_j \oplus T_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)]
\end{aligned}$$

We evaluate the first sum. For given $1 \leq i < j \leq m$, if $S_i \neq S_j$ and $T_i \neq T_j$, then the probability is smaller than $\frac{2}{2^n}$. If $S_i = S_j$ or $T_i = T_j$, it is easy to see that it is 0. Therefore

$$\sum_{i<j} \Pr[C_i \oplus T_i = C_j \oplus T_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] \leq \frac{m(m-1)}{2} \cdot \frac{2}{2^n} \leq \frac{m^2}{2^n} \quad (15)$$

The second sum can be bounded similarly.

We now consider the third sum. Let $1 \leq i, j \leq m$. As there are $2^n - m + \sigma$ possible values of C_j satisfying $\mathbf{C} \sim \mathbf{S}$ and $(\mathcal{C}_3)$, we obtain $\Pr[C_j = A_i \oplus B_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] = \frac{1}{2^n-m+\sigma} \leq \frac{2}{2^n}$. Therefore

$$\sum_{i,j} \Pr[A_i = B_j \oplus C_j | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] \leq m^2 \cdot \frac{2}{2^n} \quad (16)$$

The fourth sum can be bounded similarly.

Putting these inequalities together, we finally get

$$\Pr[\mathbf{C} \oplus \mathbf{T} \in I^{\neq} \wedge \mathbf{B} \oplus \mathbf{C} \in I^{\neq} \wedge (\mathcal{C}_2) | \mathbf{C} \sim \mathbf{S} \wedge (\mathcal{C}_3)] \geq 1 - \frac{6m^2}{2^n} \quad (17)$$

The probabilities in (14) are easy to evaluate. Thus (14) is lower bounded by:

$$\frac{\left[\frac{(2^n-2m)!}{2^{2^{n-1}-m} \cdot (2^{n-1}-m)!}\right]^2 \cdot \left[\frac{(2^n-m+\sigma)!}{2^{2^{n-1}-m+\sigma} \cdot (2^{n-1}-m+\sigma)!}\right]}{\left[\frac{2^n!}{2^{2^{n-1}} \cdot (2^{n-1})!}\right]^3} \cdot (1 - \frac{6m^2}{2^n}) \quad (18)$$

which is greater or equal than

$$2^{3m-\sigma} \cdot \left(\frac{2^{n-1}-m}{2^n-m}\right)^{3m-\sigma} \cdot \frac{1}{2nm} \geq \left(1 - \sum_{k=1}^{\infty} \frac{m^k}{2^{nk}}\right)^{3m} \cdot \frac{1}{2nm} \quad (19)$$

It remains to evaluate

$$\sum_{\substack{\mathbf{A},\mathbf{B}\in I^{\neq} \\ \mathbf{A},\mathbf{B} \text{ satisfy } (\mathcal{C}1)}} \Pr[(c_1(\mathbf{L}) \oplus \mathbf{R} = \mathbf{A}) \wedge (c_2(\mathbf{R}) \oplus \mathbf{A} = \mathbf{B})] \tag{20}$$

which is equal to

$$\begin{aligned}
&\Pr[c_1(\mathbf{L}) \oplus \mathbf{R} \in I^{\neq} \wedge c_1(\mathbf{L}) \oplus c_2(\mathbf{R}) \oplus \mathbf{R} \in I^{\neq} \\
&\quad \wedge \nexists i,j : c_1(L_i) = L_j \wedge \nexists i,j : c_2(R_i) = R_j] \\
&\geq 1 - \sum_{i<j} \Pr[c_1(L_i) \oplus c_1(L_j) = R_i \oplus R_j] \\
&\quad - \sum_{i<j} \Pr[c_1(L_i) \oplus c_2(R_i) \oplus R_i = c_1(L_j) \oplus c_2(R_j) \oplus R_j] \\
&\quad - \sum_{i<j} \Pr[c_1(L_i) = L_j] - \sum_{i<j} \Pr[c_2(R_i) = R_j]
\end{aligned}$$

Let $1 \leq i < j \leq m$. $\Pr[c_1(L_i) \oplus c_1(L_j) = R_i \oplus R_j]$ is easy to evaluate. If $R_i \oplus R_j = 0$, then $L_i \neq L_j$ and the probability is 0. If $R_i \oplus R_j \neq 0$, we can apply lemma 3. Thus in any case

$$\Pr[c_1(L_i) \oplus c_1(L_j) = R_i \oplus R_j] \leq 4/2^n \tag{21}$$

For shortness, let us denote $Z(R_i, R_j) := c_2(R_i) \oplus c_2(R_j) \oplus R_i \oplus R_j$. The terms of the second sum can be written:

$$\begin{aligned}
&\Pr[c_1(L_i) \oplus c_1(L_j) = Z(R_i, R_j)] \\
&= \Pr[c_1(L_i) \oplus c_1(L_j) = Z(R_i, R_j) | Z(R_i, R_j) = 0] \cdot \Pr[Z(R_i, R_j) = 0] \\
&\quad + \Pr[c_1(L_i) \oplus c_1(L_j) = Z(R_i, R_j) | Z(R_i, R_j) \neq 0] \cdot \Pr[Z(R_i, R_j) \neq 0] \\
&= \Pr[c_1(L_i) \oplus c_1(L_j) = 0] \cdot \Pr[Z(R_i, R_j) = 0] \\
&\quad + \Pr[c_1(L_i) \oplus c_1(L_j) = Z(R_i, R_j) | Z(R_i, R_j) \neq 0] \cdot \Pr[Z(R_i, R_j) \neq 0]
\end{aligned}$$

If $R_i = R_j$ then $L_i \neq L_j$ and the first term is 0. Else by lemma 3 it is not greater than $4/2^n$. Using lemma 3 again, the second term is also not greater than $4/2^n$. The conclusion is that

$$\Pr[c_1(L_i) \oplus c_2(R_i) \oplus R_i = c_1(L_j) \oplus c_2(R_j) \oplus R_j] \leq \frac{8}{2^n} \tag{22}$$

Finally using (21) and (22), (20) is greater or equal than

$$1 - \frac{m(m-1)}{2} \cdot \frac{4}{2^n} - \frac{m(m-1)}{2} \cdot \frac{8}{2^n} - 2 \cdot \frac{m(m-1)}{2 \cdot (2^n - 1)} \geq 1 - \frac{8m^2}{2^n} \tag{23}$$

Multiplying (19) and (23), we get

$$\mathcal{P}^{(\mathbf{L},\mathbf{R})}_{(\mathbf{S},\mathbf{T})} \geq \left(1 - \sum_{k=1}^{\infty} \frac{m^k}{2^{nk}}\right)^{3m} \cdot \frac{1}{2nm} \cdot \left(1 - \frac{8m^2}{2^n}\right) \tag{24}$$

which is greater or equal than (see proof of lemma 1)

$$\left(1 - \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{8m^2}{2^n}\right)^k\right) \cdot \left(1 - \frac{8m^2}{2^n}\right) \cdot \frac{1}{2nm} = \left(1 - \frac{12m^2}{2^n}\right) \cdot \frac{1}{2nm} \tag{25}$$

8 Conclusion and Open Problems.

In this paper we showed that replacing the inner permutations of a Misty structure by involutions without fixed point, without changing the number of rounds, did not significantly affect the previously known security bounds.

Several open problems remain: first, one could wonder whether the hypothesis "without fixed point" is important. Intuitively it is clearly not, as taking the inner permutations from a (much) bigger set increases the variety of functions one can generate, and hence the difficulty to distinguish them from perfect random functions.

Also, it is an open question whether in some cases involutions achieve significantly weaker security bounds than permutations. It should be interesting to consider involutions as inner functions of structures different from the Misty ones.

Finally, being able to do security proofs when the inner functions are even more specific (i.e. drawn from a smaller set) than involutions without fixed point would be nice, as it could maybe pave the way to security proofs on structures closer to real-life block ciphers.

References

1. P.S.L.M. Barreto and V. Rijmen. The Khazad Legacy-Level Block Cipher. Submitted as a NESSIE Candidate Algorithm. Available at `http://www.cryptonessie.org`.
2. T. Iwata, T. Yoshino, and K. Kurosawa. Non-cryptographic Primitive for Pseudorandom Permutation. In Joan Daemen and Vincent Rijmen, editors, *Fast Software Encryption, 9th International Workshop, FSE 2002, Leuven, Belgium, February 4-6, 2002*, volume 2365 of *Lecture Notes in Computer Science*, pages 149–163. Springer-Verlag, 2002.

3. T. Iwata, T. Yoshino, T. Yuasa, and K. Kurosawa. Round Security and Super-Pseudorandomness of MISTY Type Structure. In Mitsuru Matsui, editor, *Fast Software Encryption, 8th International Workshop, FSE 2001, Yokohama, Japan, April 2-4, 2001*, volume 2355 of *Lecture Notes in Computer Science*, pages 233–247. Springer-Verlag, 2002.
4. M. Luby and C. Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal on Computing*, 17(2):373–386, 1988.
5. S. Lucks. Faster Luby-Rackoff Ciphers. In Dieter Gollmann, editor, *Fast Software Encryption, Cambridge, UK, February 21-23, 1996*, volume 1039 of *Lecture Notes in Computer Science*, pages 189–203. Springer-Verlag, 1996.
6. M. Matsui. New Block Encryption Algorithm MISTY. In Eli Biham, editor, *Fast Software Encryption, 4th International Workshop, FSE '97, Haifa, Israel, January 20-22, 1997*, volume 1267 of *Lecture Notes in Computer Science*, pages 54–68. Springer-Verlag, 1997.
7. M. Minier. *Preuves d'Analyse et de Sécurité en Cryptologie à Clé Secrète*. PhD thesis, LACO, Université de Limoges, September 2002.
8. M. Minier and H. Gilbert. New Results on the Pseudorandomness of Some Blockcipher Constructions. In Mitsuru Matsui, editor, *Fast Software Encryption, 8th International Workshop, FSE 2001, Yokohama, Japan, April 2-4, 2001*, volume 2355 of *Lecture Notes in Computer Science*, pages 248–266. Springer-Verlag, 2002.
9. S. Moriai and S. Vaudenay. On the Pseudorandomness of Top-Level Schemes of Block Ciphers. In Tatsuaki Okamoto, editor, *Advances in Cryptology - ASIACRYPT 2000, Kyoto, Japan, December 3-7, 2000*, volume 1976 of *Lecture Notes in Computer Science*, pages 289–302. Springer-Verlag, 2000.
10. M. Naor and O. Reingold. On the Construction of Pseudorandom Permutations: Luby-Rackoff Revisited. *Journal of Cryptology*, 12(1):29–66, 1999.
11. J. Patarin. *Etude des Générateurs de Permutations Basés sur le Schéma du DES*. PhD thesis, Université Paris VI, November 1991.
12. J. Patarin. How to Construct Pseudorandom and Super Pseudorandom Permutations from one Single Pseudorandom Function. In Rainer A. Rueppel, editor, *Advances in Cryptology - EUROCRYPT '92, Balatonfüred, Hungary, May 24-28, 1992*, volume 658 of *Lecture Notes in Computer Science*, pages 256–266. Springer-Verlag, 1993.
13. J. Patarin. About Feistel Schemes with Six (or More) Rounds. In Serge Vaudenay, editor, *Fast Software Encryption, Paris, France, March 23-25, 1998*, volume 1372 of *Lecture Notes in Computer Science*, pages 103–121. Springer-Verlag, 1998.
14. J. Patarin. Generic Attacks on Feistel Schemes. In Colin Boyd, editor, *Advances in Cryptology - ASIACRYPT 2001, Gold Coast, Australia, December 9-13, 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 222–238. Springer-Verlag, 2001.
15. J. Patarin. Luby-Rackoff: 7 Rounds Are Enough for $2^{n(1-\epsilon)}$ Security. In Dan Boneh, editor, *Advances in Cryptology - CRYPTO 2003, Santa Barbara, USA, August 17-21, 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 513–529. Springer-Verlag, 2003.
16. Z. Ramzan and L. Reyzin. On the Round Security of Symmetric-Key Cryptographic Primitives. In Mihir Bellare, editor, *Advances in Cryptology - CRYPTO 2000, Santa Barbara, USA, August 20-24, 2000*, volume 1880 of *Lecture Notes in Computer Science*, pages 376–393. Springer-Verlag, 2000.
17. K. Sakurai and Y. Zheng. On Non-Pseudorandomness from Block Ciphers with Provable Immunity Against Linear Cryptanalysis. *IEICE Trans. Fundamentals*, E80-A(1), January 1997.

18. F.-X. Standaert, G. Piret, G. Rouvroy, J.-J. Quisquater, and J.-D. Legat. ICEBERG : an Involutional Cipher Efficient for Block Encryption on Reconfigurable Hardware. In Bimal K. Roy and Willi Meier, editors, *Fast Software Encryption, 11th International Workshop, FSE 2004, Delhi, India, February 5-7, 2004*, volume 3017 of *Lecture Notes in Computer Science*, pages 279–299. Springer-Verlag, 2004.
19. S. Vaudenay. On the Lai-Massey Scheme. In Kwok-Yan Lam, Eiji Okamoto, and Chaoping Xing, editors, *Advances in Cryptology - ASIACRYPT '99, Singapore, November 14-18, 1999*, volume 1716 of *Lecture Notes in Computer Science*, pages 8–19. Springer-Verlag, 1999.