Improving Immunity of Feistel Ciphers against Differential Cryptanalysis by using Multiple MDS Matrices

Taizo Shirai* and Kyoji Shibutani

Ubiquitous Technology Laboratories, Sony Corporation
7-35 Kitashinagawa 6-chome, Shinagawa-ku, Tokyo, 141-0001 Japan
{Taizo.Shirai, Kyoji.Shibutani}@jp.sony.com

Abstract. A practical measure to estimate the immunity of block ciphers against differential and linear attacks consists of finding the minimum number of active S-Boxes, or a lower bound for this minimum number. The evaluation result of lower bounds of differentially active S-boxes of AES, Camellia (without FL/FL^{-1}) and Feistel ciphers with an MDS based matrix of branch number 9, showed that the percentage of active S-boxes in Feistel ciphers is lower than in AES. The cause is a *difference cancellation* property which can occur at the XOR operation in the Feistel structure. In this paper we propose a new design strategy to avoid such difference cancellation by employing multiple MDS based matrices in the diffusion layer of the F-function. The effectiveness of the proposed method is confirmed by an experimental result showing that the percentage of active S-boxes of the newly designed Feistel cipher becomes the same as for the AES.

Keywords. MDS, Feistel cipher, active S-boxes, multiple MDS design

1 Introduction

Throughout recent cryptographic primitive selection projects, such as AES, NESSIE and CRYPTREC projects, many types of symmetric key block ciphers have been selected for widely practical uses [16–18]. A highly regarded design strategy in a lot of well-known symmetric-key block ciphers consists in employing small non-linear functions (S-box), and designing a linear diffusion layer to achieve a high value of the minimum number of active S-boxes [2, 4, 17, 18].

If the diffusion layer guarantees a sufficient minimum number of differentially active S-boxes, and the S-boxes have low maximum differential probability, the resistance against differential attacks will be strong enough. Let a be the lower bound on the minimum number of active S-boxes, and DP_{max} be the maximum differential probability (MDP) of S-boxes. It is guaranteed that there is no differential path whose differential characteristic probability (DCP) is higher than

* The first author was a guest researcher at Katholieke Universiteit Leuven, Dept. ESAT/SCD-COSIC from 2003 to 2004.

$(DP_{max})^a$. For instance, in the case of a 128-bit block cipher using 8-bit bijective S-boxes with $DP_{max} = 2^{-6}$, the necessary condition to rule out any path with $DCP > 2^{-128}$ is that a should be at least 22. In order to determine the appropriate number of rounds of a fast and secure cipher, it is thus essential to have an accurate estimation of the lower bound a [1, 4]. Regarding this problem, finding an optimal linear diffusion is one of the research topics included in the future research agenda of cryptology proposed by STORK project in EU [19].

Comparing the minimum number of active S-boxes of two well-known ciphers, AES and Camellia without FL/FL^{-1} (denoted by Camellia*), it is shown that the ratio of the minimum number of active S-boxes to the total number of S-boxes for Camellia* is lower than for AES. Even if the diffusion matrix of Camellia* is replaced by a 8×8 MDS based matrix with branch number 9 (which is called a MDS-Feistel cipher), the ratio won't increase significantly and there is an apparent gap between these Feistel ciphers (Camellia* and MDS-Feistel) and a SPN cipher AES.

We found that the low percentage of active S-boxes in a MDS-Feistel structure is due to a *difference cancellation* which always occurs in the differential path that realizes the minimum number of active S-boxes. In such a case, the output difference of the F-function in the i-th round will be canceled completely by the output difference of $i + 2j$-th round ($j > 0$). It is obvious that one of the conditions for difference cancellations is employing an unique diffusion matrix for all F-functions.

In this work, we propose a new design strategy to avoid the difference cancellation in Feistel ciphers with SP-type F-function. We call this new strategy *multiple MDS Feistel structure design*. The basic principle of this design is as follows. Let $2r, m$ be the round number of the Feistel structure and the number of S-boxes in the F-function, respectively. We then employ $q(< r)$ $m \times m$ MDS matrices. Furthermore they are chosen that any m columns in these q MDS matrices also satisfy the MDS property. Then, at first these MDS matrices are allocated in the odd-round F-functions, then they are allocated in the even-round F-functions again keeping the involution property. This construction removes chances of difference cancellation within consecutive $2q + 1$ rounds.

We will also show an evaluation result that confirms the effectiveness of the new design, which shows that our new design strategy makes the Feistel cipher achieve a high ratio for the minimum number of active S-boxes. The new design has a ratio that is at the same level as AES.

Our results open a way to design faster Feistel ciphers keeping its advantage that the same implementation can be used for encryption and decryption except the order of subkeys.

This paper is organized as follows: In Sect. 2, we describe some definitions used in this paper. In Sect. 3, we compare the minimum number of active S-boxes of various ciphers. In Sect. 4, we explain how difference cancellation occurs. In Sect. 5, we propose our new design strategy, *multiple MDS Feistel structure design*. In Sect. 6, we investigate the effect of the multiple MDS Feistel structure design. Finally in Sect. 7, we discuss the new design and future research.

2 Preliminaries

In this section, we state some definitions and notions that are used in the rest of this paper.

Definition 1. active S-box
An S-box which has non-zero input difference is called **active S-box***.*

Definition 2. χ function
For any difference $\Delta X \in GF(2^n)$, a function $\chi : GF(2^n) \to \{0,1\}$ is defined as follows:

$$\chi(\Delta X) = \begin{cases} 0 \; if \;\; \Delta X = \mathbf{0} \\ 1 \; if \;\; \Delta X \neq \mathbf{0} \end{cases}$$

For any differential vector $\Delta X = (\Delta X[1], \Delta X[2], \ldots, \Delta X[m]) \in GF(2^n)^m$, the truncated difference $\delta X \in \{0,1\}^m$ is defined as

$$\delta X = \chi(\Delta X) = (\chi(\Delta X[1]), \chi(\Delta X[2]), \ldots, \chi(\Delta X[m]))$$

Definition 3. (truncated) Hamming weight of vector in $GF(2^n)^m$
Let $v = (v_1, v_2, \ldots, v_m) \in GF(2^n)^m$. the Hamming weight of a vector v is defined as follows:

$$w_h(v) = \sharp\{v_i | v_i \neq 0, 1 \leq i \leq m\}.$$

Theorem 1. *[7] A $[k+m, k, d]$ linear code with generator matrix $G = [I_{k\times k} M_{k\times m}]$, is MDS iff every square submatrix (formed from any i rows and any i columns, for any $i = 1, 2, ..., min\{k, m\}$) of $M_{k\times m}$ is nonsingular.*

From the above theorem, we call a matrix M is a MDS matrix if every square submatrix is nonsingular.

Definition 4. Branch Number
Let $v = (v_1, v_2, \ldots, v_m) \in GF(2^n)^m$. The branch number $\mathcal{B}$ of a linear mapping $\theta : GF(2^n)^m \to GF(2^n)^m$ is defined as:

$$\mathcal{B}(\theta) = \min_{v \neq 0}\{w_h(v) + w_h(\theta(v))\}.$$

If M is a $m \times m$ MDS matrix and $\theta : x \to Mx$, then $\mathcal{B}(\theta) = m + 1$.

Definition 5. Feistel structure using SP-type F-function
A SP-type F-function is defined as the following: Let n be a bit width of bijective S-boxes, and m be a number of S-boxes employed in a F-function. In the i-th round F-function, (1) mn bit round key $k_i \in GF(2^n)^m$ and data $x_i \in GF(2^n)^m$ are XORed: $w_i = x_i \oplus k_i$. (2)w_i is split into m pieces of n-bit data, then each n-bit data is input to a corresponding S-box. (3) The output values of S-boxes regarded as $z_i \in GF(2^n)^m$ are transformed by an $m \times m$ matrix M over $GF(2^n)$: $y_i = Mz_i$.

A Feistel structure using SP-type F-function is shown in Fig. 1.

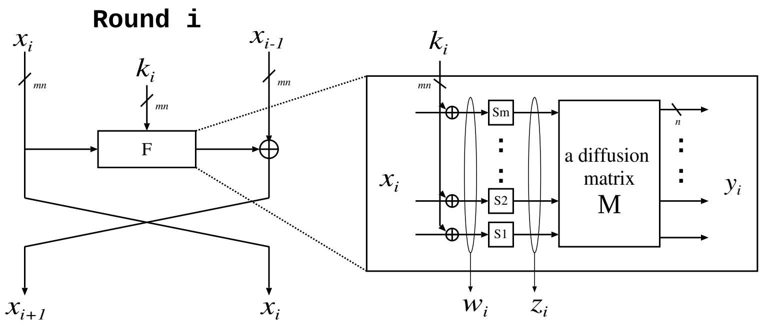


Fig. 1. The general model of a SP-type F-function

3 Comparison of the Minimum Number of Active S-Boxes

At first in this paper, we compare the lower bound of the minimum number of active S-boxes of 3 typical ciphers: AES, Camellia without FL/FL^{-1} (we call it Camellia*) and a Feistel cipher using a 8×8 MDS matrix with branch number 9 (we call it MDS-Feistel cipher). Note that we assumed that the MDS-Feistel uses eight 8-bit bijective S-boxes in the F-function like Camellia*, therefore the block sizes of these block ciphers are all 128-bit.

The lower bound estimation for these ciphers have been obtained as follows.

- AES: The wide trail strategy guarantees $\mathcal{B}^2 = 25$ active S-boxes in 4 consecutive rounds. The lower bound is obtained by using Matsui's truncated path search technique which is slightly modified to analyze AES [2, 8, 9, 12]. Let $a(r)$ be the minimum number of active S-boxes for r rounds, then the conjectured $a(r)$ from the estimation is $a(0) = 0, a(1) = 1, a(2) = 5, a(3) = 9$, and then $a(r) = a(r-4) + 25$ for $(r \geq 4)$.
- Camellia*: The lower bound is obtained from Shirai et al.'s result. They used an improved estimation method based on Matsui's technique which was also used by the designers' evaluation of Camellia [2, 8, 9, 14]. The improved method discards algebraic contradiction in difference paths [12, 13].
- MDS-Feistel: The lower bound is obtained by also using Matsui's truncated path search technique which is slightly modified to analyze the MDS-Feistel's round function [2, 8, 9, 12]. Shimizu has also shown a similar but limited result for the lower bound by using a method not based on Matsui's approach, and he has conjectured an equation $a(r) = \lfloor r/4 \rfloor (\mathcal{B}+1) + (r \mod 4) - 1$ [11]. We confirmed that our result matches the Shimizu's conjectured equation.

Table 1 shows the lower bound on the number of active S-boxes for r-round ciphers, and the ratio of active S-boxes to all S-boxes in the r-round cipher. Fig. 2 shows a graph of the ratios of active S-boxes to all S-boxes.

Round	AES	(ratio)	Camellia*	(ratio)	MDS($\mathcal{B}=9$)	(ratio)
1	1	6.3%	0	0.0%	0	0.0%
2	5	15.6%	1	6.3%	1	6.3%
3	9	18.8%	2	8.3%	2	8.3%
4	25	39.1%	7	21.9%	9	28.1%
5	26	32.5%	9	22.5%	10	25.0%
6	30	31.3%	12	25.0%	11	22.9%
7	34	30.1%	14	25.0%	12	21.4%
8	50	39.1%	16	25.0%	19	29.7%
9	51	35.4%	20	27.8%	20	27.7%
10	55	34.4%	22	23.8%	21	26.3%
11	59	33.5%	24	27.5%	22	25.0%
12	75	39.1%	-	-	29	30.2%
13	76	36.5%	-	-	30	28.8%
14	80	36.5%	-	-	31	27.7%
15	84	35.0%	-	-	32	26.7%
16	100	39.1%	-	-	39	30.5%
∞	-	39.1%	-	-	-	34.4%

Table 1. the lowerbound of the minimum number of active S-boxes

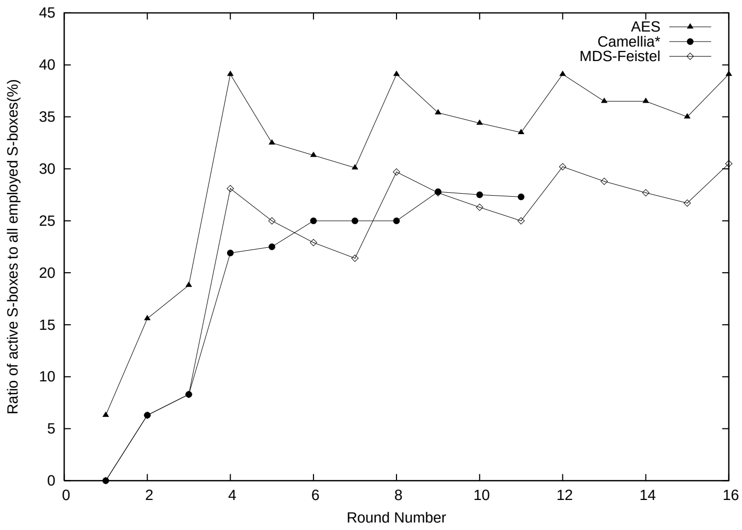


Fig. 2. The percentage of active S-boxes for AES, Camellia* and MDS-Feistel

The fact that the minimum numbers of active S-boxes are smaller for Feistel ciphers (Camellia* and MDS-Feistel) than for AES is not unexpected, because there are only half as many S-boxes in each round (8 in Feistel ciphers, 16 in AES). However, Fig. 2 shows a non-trivial fact that also the percentage of active S-boxes is lower in these Feistel ciphers than in AES. Also, we note that even with a MDS matrix of branch number 9, which is the best possible branch number for an 8×8 matrix, the construction doesn't gain significantly compared to Camellia*, which uses a non-MDS matrix of branch number 5 .

The percentage of active S-boxes indicates how many S-boxes are effectively used in all existing S-boxes for the first consecutive rounds, and can be considered as a reference of efficiency of the diffusion property for ciphers which have the same block length and the same S-box bit length.

As described in [1], if we choose 8-bit S-boxes with maximum differential probability (MDP) 2^{-6}, 22 active S-boxes is a necessary condition to rule out the existence of differential characteristics with a probability higher than 2^{-128}. In the case of AES, 22 active S-boxes are already achieved by only 4 rounds. However, Feistel ciphers require more than 11 rounds to guarantee 22 active S-boxes. Because the minimum number of active S-boxes is often taken into consideration when determining the round number of a block cipher, if more active S-boxes can be guaranteed in SP-type Feistel ciphers we may be able to design fewer rounds (it means fast) ciphers.

In the following sections, we analyze a mechanism that explains why the ratio for MDS-Feistel ciphers is low, and we propose a new design strategy that achieves more active S-boxes and thus enables us to construct Feistel structures with fewer rounds.

4 Difference Cancellation

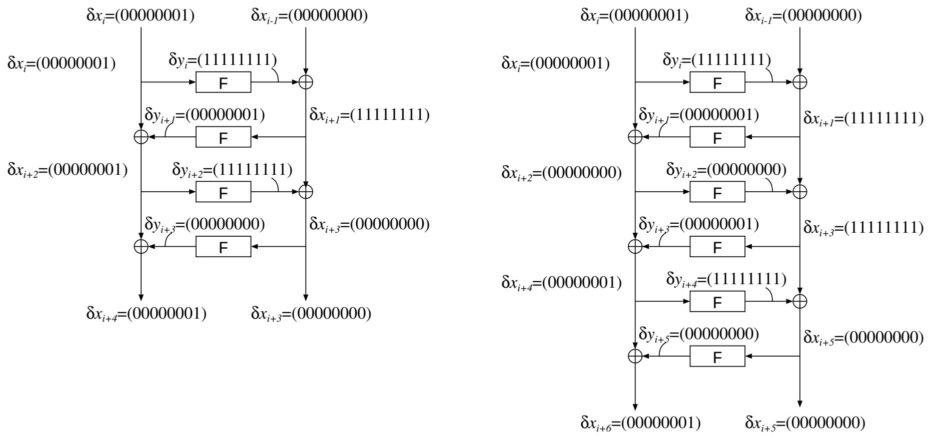
By our analysis of the MDS-Feistel cipher, we found that every path which contains the minimum number of active S-boxes includes a particular phenomenon where differences generated at certain round are canceled after some rounds at an XOR operation. We will call this phenomenon *difference cancellation*.

The left half of Fig. 3 shows an example of the 3-round difference cancellation. Differences are represented in the truncated way in which 8-bit difference data is represented as 0 or 1, depending on whether each difference is 0 or not [6]. The 3-round difference cancellation starts from the difference $\delta x_{i-1} = (00000000)$,and ends with the difference $\delta x_{i+3} = (00000000)$ again. This means that a certain difference is generated and then canceled between these two 0-differences. In this case, the full hamming weight difference $\delta y_i = (11111111)$ is canceled by $\delta y_{i+2} = (11111111)$ at once. Consequently there's no active S-boxes in $i+3$-round.

Similarly, 5-round difference cancellation, which is shown in the right half of Fig. 3, have the form $\delta x_i = \delta x_{i+4} = (00000001), \delta x_{i+2} = (00000000)$, and output differences of both active S-boxes in the i-th and $i+4$-th round are equal.

In both cases, a truncated difference (11111111) generated by one active S-box is canceled by a truncated differences (11111111) which is also generated by one active S-box. These difference cancellations are derived from 2 active S-boxes, and we call this type of difference cancellation *2-derived difference cancellation.*

An interesting fact is that at least one of these 3-round or 5-round 2-derived difference cancellations can be found in every differential path of more than 6-round that realizes the minimum number of active S-boxes in the MDS-Feistel cipher. Details are shown in Appendix A.



3-round Difference Cancellation

5-round Difference Cancellation

Fig. 3. Difference Cancellation

4.1 Observation on Difference Cancellation

Let $X, Y \in \{0,1\}^8$ and $X \overset{M}{\rightarrow} Y$ denotes that a truncated difference X can produce a truncated difference Y by a matrix M. Let M_{MDS} be a 8×8 MDS matrix, then the following property of the M_{MDS} contributes to occur the 2-derived difference cancellation.

$$(00000001) \overset{M_{MDS}}{\longrightarrow} (11111111) \quad and \quad (11111111) \overset{M_{MDS}}{\longrightarrow} (00000001)$$

These transitions are appeared in the above 3-round and 5-round difference cancellation several times.

More precisely, let $C_M(X,Y) = \{0,1\}$ be a function which shows the capability of connection between truncated difference X and Y defined as,

$$C_M(X,Y) = \begin{cases} 1 & if \quad X \overset{M}{\rightarrow} Y \\ 0 & else \end{cases}$$

We can observe that a 2-derived difference cancellation can occur if there exists at least one set of truncated differences X, Y where $w_h(X) = 1$ which satisfy $C_M(X, Y) \cdot C_M(Y, X) \neq 0$.

From MDS property, any $m \times m$ MDS matrix M_{MDS} holds the condition $C_{M_{MDS}}(X, Y) = 1$ for all $w_h(X) + w_h(Y) \geq m + 1$ and $w_h(X) = w_h(Y) = 0$. Otherwise $C_{M_{MDS}}(X, Y) = 0$. It is obvious that at least one set X, Y where $w_h(X) = 1$ satisfy $C_{M_{MDS}}(X, Y) * C_{M_{MDS}}(Y, X) \neq 0$, thus 2-derived difference cancellation can occur in a MDS matrix construction.

This observation explains why Camellia*'s lower bounds are not too low even though it employs a non-MDS matrix M_{Ca} of branch number 5. For any choice of X, Y where $w_h(X) = 1$, $C_{M_{Ca}}(X, Y) \cdot C_{M_{Ca}}(Y, X) = 0$ always. Thus M_{Ca} never produces the 2-derived difference cancellation, and it keeps a moderate number of active S-boxes.

However, even though 2-derived difference cancellation is avoided by choosing Camellia type matrix, if certain X, Y where $w_h(X) = 2$ satisfying $C_M(X, Y) \cdot C_M(Y, X) \neq 0$ exists, then a 4-derived difference cancellation would be a building block for a small number of active S-boxes, and a significant gain of the number of active S-boxes may not be expected.

In the next section another approach to avoid m-derived difference cancellation will be introduced by using multiple MDS matrices in a Feistel structure.

5 Multiple MDS Feistel Structure Design

5.1 Basic Strategy

Suppose that some intermediate differential data $\Delta x_{i-1} = 0$, and that the output of F-function in every 2 rounds is added to the data, (ex. $\Delta y_i, \Delta y_{i+2}, ..\Delta y_{i+2j}$). Consider a situation where the differential data Δx_{i+2j+1} become 0 after XORing the output of the F-function in the $i + 2j$-th round, caused by a difference cancellation as shown in Fig. 4.

In the difference cancellation, the following condition exists:

$$\sum_{k=0}^{j} \Delta y_{i+2k} = 0\,. \tag{1}$$

Therefore,

$$M \sum_{k=0}^{j} \Delta z_{i+2k} = 0\,. \tag{2}$$

When a nonsingular matrix M is employed, we obtain that

$$\sum_{k=0}^{j} \Delta z_{i+2k} = 0\,. \tag{3}$$

The above equation shows that a difference cancellation occurs by only 2 active S-boxes in $\Delta z_{i+2k}, (0 \leq k \leq j)$ in the minimum case, which is exactly the case of

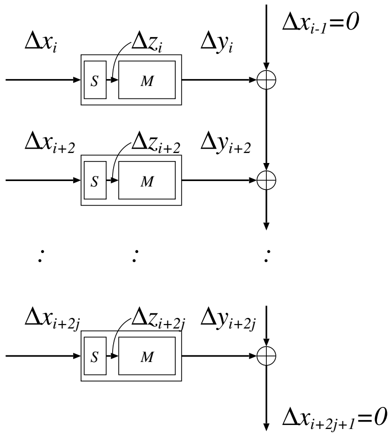


Fig. 4. Difference Cancellation

2-derived difference cancellations shown in the previous section. Now we consider a setting with multiple matrices in which a different matrix is used in each F-function. Let M_i be a diffusion matrix employed in the i-th round. Obviously, the transformation from (1) to (2) is not correct when the matrices M_i are different from each other. In such setting, we can rewrite (1) as

$$M_i \Delta z_i + M_{i+2} \Delta z_{i+2} + \ldots + M_{i+2j} \Delta z_{i+2j} = 0\,. \tag{4}$$

The above condition can be written as the product of a large $m \times m(j+1)$ matrix and a vector with $m(j+1)$ elements:

$$[M_i M_{i+2} \cdots M_{i+2j}] \begin{bmatrix} \Delta z_i \\ \Delta z_{i+2} \\ \vdots \\ \Delta z_{i+2j} \end{bmatrix} = 0\,. \tag{5}$$

If these matrices are chosen to satisfy that there is no combination of l column vectors that are dependent of each other $(2 \le l \le m)$ in the matrix, k-derived difference cancellation $(k \le l)$ would never happen in the consecutive $2j$ rounds. From this observation, we introduce a strategy to choose matrices $M_i, .., M_{i+2j}$ for which any choice of m column vectors are independent of each other in the large matrix $[M_i, .., M_{i+2j}]$.

5.2 Construction Steps

We propose a new design strategy that employs multiple MDS matrices in the Feistel network, in order to avoid an occurrence of m-derived difference cancellation in any consecutive $2q$ rounds where q is the number of employed matrices.

The construction steps are as follows. Without loss of generality, we assume the round number is $2r$.

1. Choose $q(\leq r)$ MDS matrices: $M_0, M_1, \ldots, M_{q-1}$.
2. Check that any m of qm column vectors in all M_i matrices hold the MDS property.
3. Assign matrix $M_{(i \mod q)}$ to the $2i+1$-th round $(0 \leq i < r)$.
4. Assign matrix $M_{(i \mod q)}$ to the $2r-2i$-th round $(0 \leq i < r)$ (reverse order).

In this construction, since any m columns of the large matrix $[M_i M_{i+2} \cdots M_{i+2q-2}]$ have been chosen to generate MDS which has m independent column vectors, there is no chance to generate m-derived difference cancellation in any consecutive $2q-1$ rounds. Fig. 5 shows the construction of the example setting $r = 6, q = 3, 6$ respectively.

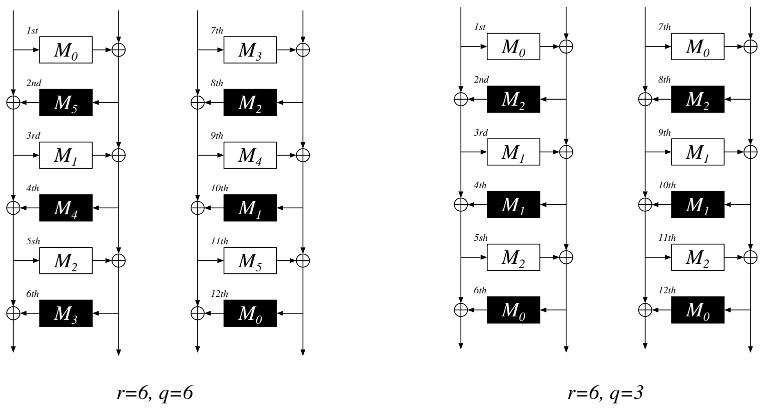


Fig. 5. Examples of the New Design $(r = 3, q = 3, 6)$

When m, n are small, we can randomly generate MDS matrices and check MDS conditions of column vectors in Step 2. However, if m, n are large, it might be difficult to search such a set of matrices. In such a case, we can use the algorithm to generate Reed-Solomon code's generation matrix. The algorithm can generate a large MDS matrix immediately, because its complexity is $O(N^3)$ where N is the dimension of the matrix [7]. Once the $qm \times qm$ MDS matrix M_L is made, any combination of m rows of M_L can be regarded as a matrix $[M_0, M_1, .., M_{q-1}]$ with proposed additional MDS property from Theorem 1. We can find a 128×128 MDS matrix on $GF(2^8)$ from $[256, 128, 129]$ extended RS codes, thus 16 MDS matrices of dimension 8 satisfying the condition of Step 2. can be found in it. We show an example of a set of such matrices in Appendix B.

6 Evaluation of the Proposed Construction

We estimated a lower bound for the number of active S-boxes of the new construction with $m = 8, r = 6$ (12-round cipher) for the number of matrices $q = 1..6$. We adopted a weight based approach in the evaluation algorithm, because the known truncated path approach which is employed in the other cipher's evaluation requires too huge memory space and time consumption.

6.1 Algorithm

The following algorithm outputs a lower bound for the number of active S-boxes of our proposed construction based on the weight based approach . Let the round number be R.

1. Set $L = \infty$.
2. For each possible combination of the weight $0, 1, .., 8$ in $\delta x_0, \delta x_1, .., \delta x_{R+1}$ (There are 9^{R+2} candidates):
   (a) For $i = 2$ to $R + 1$ do the following,
      i. For $j = 2$ to $j \leq i$, $j \leftarrow j + 2$ do the following,
         A. Check whether the given weight combination of $w_h(\delta x_{i-j}), w_h(\delta x_i)$ and the list of given weight of active S-boxes in the $w_h(\delta x_{i-j+1})$, $w_h(\delta x_{i-j+3})$,.., $w_h(\delta x_{i-1})$ are possible or not in the weight context of the given MDS property.
         B. If the check passed then continue the loop, else exit the loop.
   (b) If all checks passed, count the total number of active S-boxes A in the path. If $A < L$, set $L = A$.
3. Output L as the lower bound of the minimum number of active S-boxes for the round R.

The description to check the possibility of a weight distribution in Step A is described in Appendix C.

We note that this algorithm can be speeded up for R-round evaluation by using the result of $R - 1$-round evaluation recursively. This technique can be seen in Matsui's path search method [8].

6.2 Result

Table 2 shows the result of the lower bound of the minimum number of active S-boxes for four types of 12-round multiple-MDS Feistel ciphers, the cases of $m = 8, r = 6, q = 1, 2, 3, 6$. The graph of the ratio of active S-boxes to total number of S-boxes is shown in Fig. 6. It can be confirmed that the result of the case $q = 1$ is the same as the result of the MDS-Feistel cipher shown in Table 1. Moreover, the results shows that the lower bounds for the cases $q = 3$ and $q = 6$ are always the same.

The numbers are significantly increased in the case of $q = 2$ compared to the case $q = 1$. However there is no gain in 8 and 9 rounds. In the case of $q \geq 3$,

Round	$q=1$	(rate)	$q=2$	(rate)	$q=3$	(rate)	$q=6$	(rate)
1	0	0.0%	0	0.0%	0	0.0%	0	0.0%
2	1	6.3%	1	6.3%	1	6.3%	1	6.3%
3	2	8.3%	2	8.3%	2	8.3%	2	8.3%
4	9	28.1%	9	28.1%	9	28.1%	9	28.1%
5	10	25.0%	10	25.0%	10	25.0%	10	25.0%
6	11	22.9%	18	37.5%	18	37.5%	18	37.5%
7	12	21.4%	18	32.1%	18	32.1%	18	32.1%
8	19	29.7%	19	29.7%	20	31.3%	20	31.3%
9	20	27.7%	20	27.8%	27	37.5%	27	37.5%
10	21	26.3%	27	33.8%	28	35.0%	28	35.0%
11	22	25.0%	28	31.8%	32	36.4%	32	36.4%
12	29	30.2%	36	37.5%	36	37.5%	36	37.5%
avrg.(4-12)		26.3%		31.5%		33.4%		33.4%

Table 2. Result of Evaluation

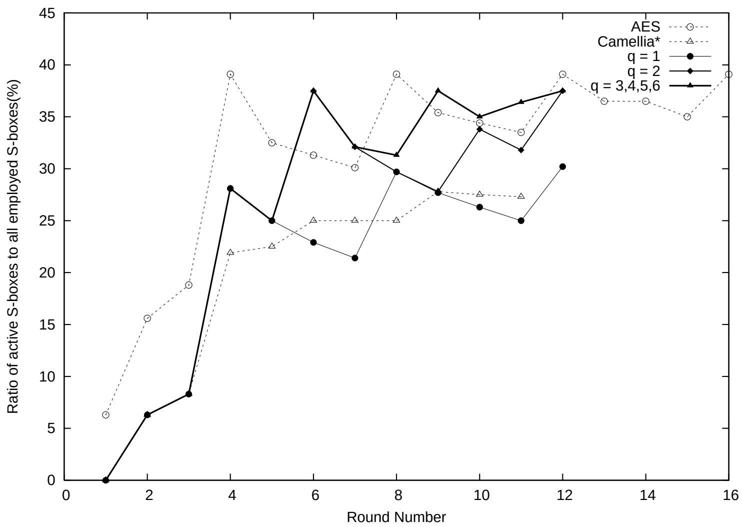


Fig. 6. The percentage of active S-boxes for AES, Camellia* and multiple-MDS Feistel

the lower bound is even higher than for the case $q = 2$ when we have more than 8 rounds, and more than 22 active S-boxes are guaranteed in 9 rounds. The ratio of the new design successfully came to the level of AES after more than 6 rounds. These results show that the design with triple MDS matrices has enough advantages over single MDS matrix design. Also, our experiment indicates that not so many MDS matrices seem to be required to get a benefit from the proposed design.

7 Discussion

7.1 Implementation Aspects

This new construction requires an additional implementation cost because it employs multiple matrices. Multiple diffusion matrices require additional gate size in hardware and lookup tables in memory in software implementation. However the speed impact in hardware is expected to be negligible because only switching circuits for matrices will be added. Detailed observations on hardware implementation of many types of SP-type Feistel networks can be found in [15].

If all lookup tables can be stored in the fastest cache memory, not much time cost would be expected. If b matrices of dimension 8 on $GF(2^8)$ are employed in the 128-bit block setting, the cipher requires $16b$ KB lookup tables at maximum in which the size of each entry is 64-bit. Since some recent 64-bit CPUs have 64KB first cache memory for data, 48 KB lookup table required by 3 matrices would be acceptable. In such a setting, it is estimated that only $8R$ table lookups and $9R$ XOR operations are required to finish R round calculation without a key scheduling procedure.

7.2 Future Research

Though we only discussed the immunity against differential attacks throughout this paper, we can directly extend the result to the linear attack if we construct a PS-type F-function whose order of S-box and diffusion layer is exchanged from SP-type F-function[5]. This is due to the dual property of differential and linear masks [5, 10]. However, it is not clear so far that the immunity has been gained for the linear attack if a cipher is designed to have immunity against the differential attack based on our strategy. This theoretical explanation should be included in the topic of future research.

Our evaluation method adopted a simple weight based approach to estimate lower bounds of the proposed designs. Since the approach achieved an algorithm with feasible time and memory space at the expense of information of truncated differential form, a more detailed algorithm may produce tighter lower bounds. It is considered an important research topic to develop a new algorithm that counts lower bounds of the minimum number of active S-boxes more strictly for the new design.

Acknowledgment

The authors would like to thank Bart Preneel, Alex Biryukov, Christophe De Cannière and Joseph Lano of COSIC, K.U.Leuven for their helpful discussions and suggestions. Furthermore, we thank the anonymous reviewers for helpful feedback.

References

1. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita, "Specification of Camellia - a 128-bit Block Cipher," Primitive submitted to NESSIE by NTT and Mitsubishi, 2000, See also http://info.isl.ntt.co.jp/camellia
2. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita, "Camellia: A 128-Bit Block Cipher Suitable for Multiple Platforms - Design and Analysis," Selected Area in Cryptography, SAC 2000, LNCS 2012, pp.39-56, 2000.
3. E. Biham and A. Shamir," Differential Cryptanalysis of DES-like Cryptosystems," CRYPTO '90, LNCS 537, pp.2-21, 1991.
4. J. Daemen and V. Rijmen, "The Design of Rijndael: AES - The Advanced Encryption Standard," Springer-Verlag, 2002
5. M. Kanda, "Practical Security Evaluation against Differential and Linear Cryptanalysis for Feistel Ciphers with SPN Round Function," Selected Areas in Cryptography, SAC 2000, LNCS 2012, pp. 324-338, 2000.
6. L.R.Knudsen, "Truncated and Higher Order Differentials," Fast Software Encryption - Second International Workshop, LNCS 1008, pp.196-211, 1995.
7. R. Lidl and H. Niederreiter, "Finite Fields," Encyclopedia of mathematics and its applications 20, Cambridge Univ. Press, 1997.
8. M.Matsui, "Differential Path Search of the Block Cipher E2," Technical Report ISEC99-19, IEICE, 1999.(written in Japanese)
9. M. Matsui and T. Tokita,"Cryptanalysis of Reduced Version of the Block Cipher E2," Fast Software Encryption, FSE'99, LNCS 1636, 1999.
10. M. Matsui, "Linear Cryptanalysis of the Data Encryption Standard," EUROCRYPT '93, LNCS 765, pp.386-397, 1994.
11. H. Shimizu, "On the security of Feistel cipher with SP-type F function," 7A-3, SCIS 2001, 2001. (written in Japanese)
12. T. Shirai, S. Kanamaru and G. Abe, "Improved Upper Bounds of Differential and Linear Characteristic Probability for Camellia", Fast Software Encryption, FSE2002, LNCS 2365, pp.128-142, 2002.
13. T. Shirai, "Differential, Linear, Boomerang and Rectangle Cryptanalysis of Reduced-Round Camellia," preproceedings of Third NESSIE Workshop, 2002.
14. S. Moriai, M. Sugita, K. Aoki and M. Kanda, "Security of E2 against Truncated Differential Cryptanalysis," Selected Areas in Cryptography, SAC'99, LNCS 1758, pp.106-117, 2000.
15. L. Xiao and H. M. Heys,"Hardware Performance Characterization of Block Cipher Structures," Topics in Cryptology - CT-RSA 2003, The Cryptographers' Track at the RSA Conference 2003, Proceedings pp.176-192, 2003
16. National Institute of Standards and Technology, Advanced Encryption Standard, FIPS 197, 2001.
17. NESSIE project - New European Schemes for Signatures, Integrity, and Encryption, http://www.cryptonessie.org

18. CRYPTREC project, http://www.ipa.go.jp/security/enc/CRYPTREC/.
19. STORK project - Strategic Roadmap for Crypto, Public Document, D6: Open problems in Cryptology version 2.1, chapter 4, 2003, available at http://www.stork.eu.org/

Appendix A

All the minimum differentially active S-boxes paths for more than 5 rounds of MDS-Feistel cipher using an 8×8 MDS matrix can be represented by only the following eight types of differential paths as building blocks in Fig. 7.

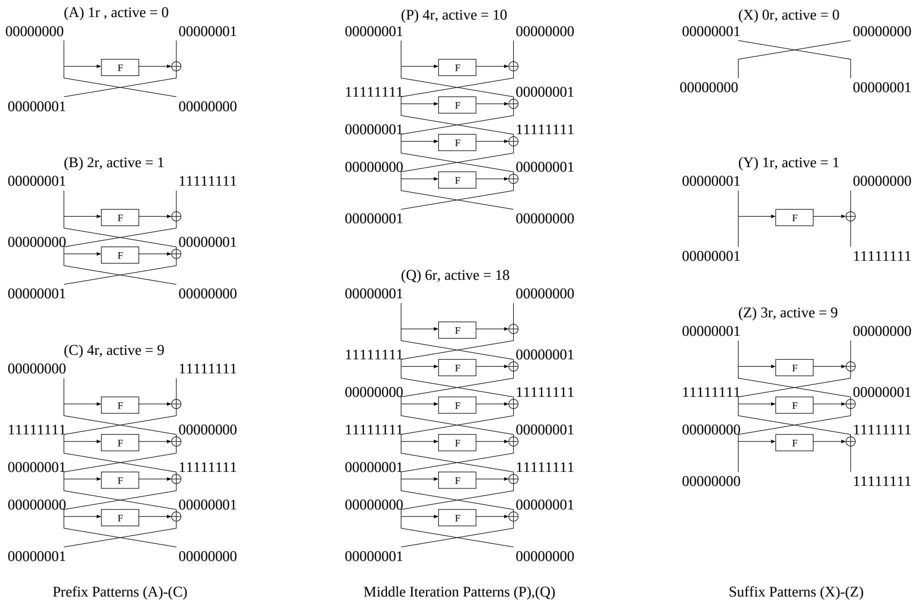


Fig. 7. Building Blocks based on $\delta A = (00000001)$

Pattern (A),(B) and (C) are prefix patterns which appear only at the beginning of differential paths, and pattern (X), (Y) and (Z) are suffix patterns which appear only at the end of differential paths. Pattern (P) and (Q) are middle iteration patterns which appear at the middle of differential paths and are sometimes iterated more than once depending on the total number of rounds. (P) and (Q) are respectively shown as 3-round and 5-round differential cancellations in Sect. 4.

Each pattern in Fig. 7 shows a representative path using truncated difference (00000001), and each pattern also contains 7 other different paths by replacing (00000001) with one of $(00000010), (00000100), .., (10000000)$.

Table 3 shows the search results for 5-round to 20-round differential paths of MDS-Feistel. The *Patterns* field shows the path patterns expressed by their building blocks. It means that any resulting path can be expressed by one of the path patterns in the corresponding field. We can see that there is a 4-round regularity. The last three rows shows the regularity in a generalized form. From this experimental result, any differential path with a minimum number of active

S-boxes for more than 6-rounds contains at least one (P) or (Q). This is the reason why difference cancellation should be avoided in order to gain more active S-boxes, as described in Sect. 4.

R.	1	2	3	4
M.A.	0	1	2	9
R	5	6	7	8
M.A.	10	11	12	19
Pat.	APX BZ CY	APY BPX	BPY	APZ AQY BQX CPX
R.	9	10	11	12
M.A.	20	21	22	29
Pat.	APPX BPZ BQY CPY	APPY BPPX	BPPY	APPZ APQY, AQPY BPQX, BQPX CPPX
R.	13	14	15	16
M.A.	30	31	32	39
Pat.	APPPX BPPZ BPQY, BQPY CPPY	APPPY BPPPX	BPPPY	APPPZ APPQY, APQPY, AQPPY BPPQX, BPQPX, BQPPY CPPPX
R.	17	18	19	20
M.A.	40	31	32	39
Pat.	APPPPX BPPPZ BPPQY, BPQPY, BQPPY CPPPY	APPPPY BPPPPX	BPPPPY	APPPPZ APPPQY, APPQPY, APQPPY, AQPPPY BPPPQX, BPPQPX, BPQPPX, BQPPPX CPPPPX
R.	4n+1	4n+2	4n + 3	4n + 4
M.A.	10n	10n+1	10n+2	10n + 9
Pat. ($R. \geq 5$)	AP^nX $BP^{n-1}Z$ $B(P^{n-2}Q)Y (n > 1)$ $CP^{n-1}Y$	AP^nY BP^nX	BP^nY	AP^nZ $A(P^{n-1}Q)Y$ $B(P^{n-1}Q)X$ CP^nX

R. : Round Number
M.A. : the minimum numbers of active S-boxes
Pat. : path expressions constructed from basic patterns.
P^k: iteration of pattern P for k times
(P^kQ): all possible patterns generated from pattern P for k times and Q for once

Table 3. All path patterns of the minimum number of active S-boxes for each round

Appendix B

In this appendix, we show a set of example 8×8 MDS matrices in which any combination of any 8 columns form a MDS matrix, which was obtained from the right part of a $[256, 128, 129]$ Reed-Solomon code's generation matrix in standard form [7].

In the following example we employed a primitive polynomial $p(x) = x^8 + x^4 + x^3 + x^2 + 1$. Let α be a root of p(x), we set a parity check matrix H as:

$$H = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 \\ \alpha^{254} & \alpha^{253} & \cdots & \alpha & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ (\alpha^{254})^{126} & (\alpha^{253})^{126} & \cdots & \alpha^{126} & 1 & 0 \\ (\alpha^{254})^{127} & (\alpha^{253})^{127} & \cdots & \alpha^{127} & 1 & 0 \end{bmatrix}.$$

Then we calculated a generation matrix $G = [I_{128\times128} M_{128\times128}]$ The following 16 matrices are obtained from the first 8 rows of $M_{128\times128}$ simply by splitting every 8 columns as $[M_0 M_1 ... M_{15}]$. Each element is expressed in a hexadecimal value corresponding to a binary representation of elements in $GF(2^8)$.

$$M_0 = \begin{pmatrix} 9d & b4 & d3 & 5d & 84 & ae & ec & b9 \\ 29 & 34 & 39 & 60 & 5c & 81 & 25 & 13 \\ 67 & 6a & d2 & e3 & 4b & db & 9d & 4 \\ 8e & d7 & e6 & 1b & 8b & 9e & 3a & 91 \\ d9 & e5 & 4d & dd & c6 & 5 & f0 & ad \\ 2a & f7 & 67 & 72 & b1 & 7 & f2 & 27 \\ 42 & e6 & a0 & 4 & f1 & 4 & 7d & 8c \\ 55 & 63 & fa & 51 & c & d9 & 28 & d6 \end{pmatrix} \quad M_1 = \begin{pmatrix} b8 & f1 & 65 & ef & d2 & c3 & 7b & f4 \\ 3a & f6 & 2d & 6a & 1e & cc & 5e & a4 \\ 4a & 97 & a3 & b9 & f4 & 2b & a0 & 76 \\ 82 & 5f & a2 & c1 & bf & 30 & 69 & 2d \\ 59 & 89 & 10 & 2d & 4 & bc & fb & 5c \\ 1d & 69 & eb & 4e & c8 & b8 & b0 & 2d \\ 31 & 1b & 22 & 29 & 71 & 51 & 37 & 63 \\ e0 & 7b & b5 & 5a & f2 & 81 & cd & 81 \end{pmatrix} \quad M_2 = \begin{pmatrix} d0 & 46 & a6 & a7 & e1 & b7 & 16 & d2 \\ fd & b3 & 84 & 18 & 7a & cc & 31 & e7 \\ ca & 9b & d3 & 9c & 66 & b1 & 12 & af \\ 79 & ec & 6a & a8 & c1 & 55 & e2 & 14 \\ 56 & f8 & a0 & 79 & 3a & 4b & 13 & 27 \\ 77 & e1 & 26 & 19 & 77 & bd & 3a & f6 \\ c2 & 5 & 33 & 9c & d1 & 3 & 1e & 5 \\ 2b & fd & 5b & aa & 3a & a4 & 47 & c5 \end{pmatrix}$$

$$M_3 = \begin{pmatrix} ce & a2 & 8 & 3d & c2 & c4 & c0 & a6 \\ 9a & e1 & 65 & f6 & 5f & b5 & 5e & 2 \\ b0 & 7 & 6 & 6f & bb & 1f & 8 & 3e \\ d7 & 53 & 23 & 62 & 21 & ee & 58 & f9 \\ eb & fa & 91 & 69 & 3e & a2 & c & 14 \\ 62 & b4 & e5 & 2a & b2 & aa & b7 & d2 \\ dd & d & d3 & 18 & db & 2e & 8b & 65 \\ 7f & b8 & 7b & 70 & 2f & 44 & 8a & d5 \end{pmatrix} \quad M_4 = \begin{pmatrix} 1c & 42 & 16 & 4d & f4 & b5 & f2 & 71 \\ 1f & 92 & c4 & 36 & 92 & 21 & 7f & 50 \\ cb & 84 & 4c & cd & 1e & a9 & 4 & 4 \\ 40 & ff & d7 & cd & 40 & 24 & 3f & 6b \\ c3 & a0 & b9 & 75 & f8 & 74 & ce & 88 \\ b4 & f5 & d1 & b8 & a3 & b1 & ed & 13 \\ c6 & 7b & 56 & 7c & 6b & 7 & f & d5 \\ c4 & c2 & 3f & 4b & ca & 8d & 19 & 76 \end{pmatrix} \quad M_5 = \begin{pmatrix} 24 & 75 & 60 & ec & cf & de & 60 & 4f \\ a7 & 6b & 38 & eb & d8 & 14 & 93 & b8 \\ 16 & 21 & bb & 24 & c4 & 5c & c4 & 6a \\ 86 & 6b & 8c & 5a & d6 & f4 & f1 & 4c \\ 25 & 50 & 1a & e2 & fa & b0 & 85 & ed \\ 75 & 79 & d4 & ed & c0 & 81 & 34 & 4a \\ 2f & f0 & e7 & ae & ae & 25 & 1f & 49 \\ 2b & 6d & a2 & cb & 13 & 38 & 77 & 91 \end{pmatrix}$$

$$M_6 = \begin{pmatrix} 86 & 33 & 2e & b3 & 64 & c1 & 36 & 8a \\ b6 & aa & da & ee & 13 & 82 & 7c & e5 \\ 63 & 56 & fb & ec & c3 & d6 & e7 & bd \\ b & 38 & f9 & da & 53 & e & 15 & b0 \\ 2f & 5c & a & 12 & ca & ee & 67 & d5 \\ 4f & c4 & 70 & c & 17 & b8 & e3 & 5b \\ 19 & b8 & 4f & b9 & 2f & 5e & 9a & 6a \\ d2 & 64 & a3 & 1 & d0 & 4c & df & 4c \end{pmatrix} \quad M_7 = \begin{pmatrix} 7f & d7 & 37 & 4a & 8b & ca & f & e9 \\ fa & 1f & 16 & 7e & e5 & 2e & 64 & 15 \\ 74 & e1 & 2 & 6a & c7 & 4f & bc & fd \\ ed & ea & a8 & 12 & b7 & ad & fa & c5 \\ 1a & 16 & dc & 8a & 8d & 29 & ef & aa \\ 58 & 19 & 7c & f5 & d1 & 49 & 1e & fe \\ 69 & b2 & 53 & d4 & 14 & 47 & 2b & b1 \\ cd & 4e & f4 & 21 & c5 & 55 & 5e & fb \end{pmatrix} \quad M_8 = \begin{pmatrix} 50 & cd & 21 & ff & 88 & 97 & 8b & c \\ c & 9c & 8a & bd & d4 & 9e & 38 & a1 \\ ad & d3 & 5f & cd & 8c & a4 & 27 & 22 \\ 88 & 56 & c9 & 75 & 93 & 2f & 79 & 11 \\ 2 & 96 & 26 & df & b & 36 & b0 & da \\ 6c & ee & 8d & 46 & 2d & f0 & 6d & 2e \\ 2f & ff & 7 & 81 & 1f & d8 & 11 & b7 \\ ca & 1 & fd & 93 & c4 & af & c9 & 5c \end{pmatrix}$$

$$M_9 = \begin{pmatrix} 9c & ae & 3e & 9e & 58 & ca & c8 & 77 \\ 59 & 99 & d4 & 93 & bc & fd & 97 & 7f \\ be & de & 1b & 3b & cc & 16 & a7 & e1 \\ 26 & 6b & 39 & 16 & 6a & a6 & 75 & d3 \\ 97 & a8 & c0 & 1 & 13 & f4 & 86 & 1d \\ 97 & 56 & ba & 43 & d8 & d7 & fe & 14 \\ 4 & 22 & 13 & 40 & f5 & 4e & 91 & ab \\ 25 & 2c & d5 & 12 & 4d & b4 & 9c & 40 \end{pmatrix} \quad M_{10} = \begin{pmatrix} 51 & f4 & a9 & 82 & c8 & f7 & d9 & f6 \\ ee & 8e & 98 & bd & df & 93 & 45 & fe \\ bf & c6 & 7c & be & e7 & 7f & 62 & 9d \\ 1e & 32 & 82 & f & dd & e9 & de & e6 \\ 4b & 2b & 3c & 80 & 2f & 9 & f & 55 \\ bc & 9d & c2 & 1f & 8a & 50 & 5a & 17 \\ ad & 6f & 2d & 2c & 59 & e1 & b0 & 59 \\ 17 & 1b & 28 & 8d & fe & bb & 18 & 95 \end{pmatrix} \quad M_{11} = \begin{pmatrix} d7 & bf & 93 & 96 & 9 & ae & 2b & 49 \\ 96 & 3d & 44 & f9 & 2d & c & d6 & e6 \\ e5 & ba & b0 & 4c & 66 & aa & d8 & 22 \\ 3 & f2 & b & 99 & e2 & b3 & 9d & 4b \\ 1d & 4b & 36 & f1 & 4 & 6c & bf & 5e \\ 56 & 85 & 31 & aa & 89 & c5 & a6 & 3f \\ fd & 45 & cd & ac & a5 & 3c & 9b & b6 \\ 7e & fe & ce & ba & 1d & 8d & db & bd \end{pmatrix}$$

$$M_{12} = \begin{pmatrix} 9d & 16 & 3a & 75 & a0 & f2 & 4a & c2 \\ 60 & 1b & 81 & 75 & a2 & 6a & bb & 28 \\ 81 & be & 64 & 7 & 18 & 87 & 16 & f6 \\ ac & d2 & 4b & 19 & ed & 8e & 97 & 58 \\ 92 & ea & 18 & 9d & 8a & 7 & ac & cb \\ 74 & e3 & 79 & 44 & c & 13 & 2e & 77 \\ 7d & d7 & 1b & fc & fb & b2 & bb & df \\ e3 & 39 & 1b & 59 & a & e5 & c0 & 8c \end{pmatrix} \quad M_{13} = \begin{pmatrix} 54 & c8 & 7d & 23 & 25 & 3f & a9 & 99 \\ ca & 9 & 40 & c2 & 89 & 23 & 53 & 6d \\ 6d & 90 & 68 & 73 & fc & 73 & d1 & c9 \\ 91 & e & c9 & 7b & 7b & 91 & 37 & 4d \\ fe & b1 & 1a & 7 & 6a & 8f & 9f & 8f \\ 8c & f4 & f2 & cf & 9d & 1f & 66 & 34 \\ 1b & ca & 16 & cb & 9b & b0 & af & 99 \\ 9f & c9 & fe & 87 & f2 & ab & f7 & c1 \end{pmatrix} \quad M_{14} = \begin{pmatrix} a1 & 6d & 5 & 75 & 5 & 9c & 74 & d9 \\ 32 & ba & fc & 4a & e4 & 50 & af & c6 \\ 78 & 80 & 3f & 7a & cf & fb & ae & 5e \\ a9 & 69 & 18 & 42 & e2 & cb & c & f \\ 81 & 24 & 57 & ce & 7c & 64 & 42 & ea \\ 79 & fc & c2 & b & 28 & 31 & 5f & 11 \\ e3 & 55 & 29 & 47 & de & 2 & dc & bf \\ a8 & 3b & a9 & 6b & bb & 4c & 87 & 25 \end{pmatrix}$$

$$M_{15} = \begin{pmatrix} 49 & 7f & 18 & 34 & 55 & 8c & 7c & 4c \\ 14 & 39 & 2b & 5d & 40 & 93 & 78 & 10 \\ 32 & 85 & d & be & 60 & 8c & 8a & 61 \\ cd & 62 & e9 & c7 & 53 & 2c & 5f & 6a \\ cc & 67 & 39 & 3e & 3e & 4c & 67 & 97 \\ 6a & 41 & 46 & fa & 1e & 4d & 68 & f1 \\ 1 & 58 & e8 & b4 & fe & fa & c7 & 34 \\ 99 & a6 & c2 & fa & 33 & 80 & 9 & ee \end{pmatrix}$$

Appendix C

In our evaluation algorithm, a check procedure to judge whether a given weight distribution is possible or not was employed.

Let M_i $(1 \leq i \leq 2r)$ be the i-th round diffusion matrix, and let N_j $(1 \leq j \leq q)$ be q matrices designed by our proposed design strategy. In any combination of Δx_i and $\Delta x_{i+2j} \in GF(2^n)^m$, there is the following relation,

$$\Delta x_i + \Delta x_{i+2j} = \sum_{k=1}^{j} M_{i+2k-1} \Delta z_{i+2k-1} \tag{6}$$

In the case of $q = r$, all M_k in the above equation are guaranteed to be different because they are chosen from N_j without overlap.

Then we consider weight conditions on both sides of (6). Let W_1 be $w_h(\Delta x_i + \Delta x_{i+2j})$. We obtain the following inequality.

$$|w_h(\Delta x_i) - w_h(\Delta x_{i+2j})| \leq W_1 \leq min(m, w_h(\Delta x_i) + w_h(\Delta x_{i+2j})) \tag{7}$$

On the other hand, let W_2 be $w_h(\sum_{k=1}^{j} M_{i+2k-1} \Delta z_{i+2k-1})$. If there is at least one nonzero hamming weight in $w_h(\Delta z_j)$,

$$max(m + 1 - \sum_{k=1}^{j} w_h(\Delta z_{i+2k-1}), 0) \leq W_2 \leq m \tag{8}$$

If all hamming weight of $w_h(\Delta z_j)$ are 0, obviously $W_2 = 0$.

In the evaluation algorithm, we compare the above weight conditions W_1 and W_2 implied by a given weight distribution. In the checking procedure, it is judged false if these two weight range conditions have no overlap, because it means there is no path with such a weight distribution.

In $q < r$ and $j > q$ settings, some of the matrices N_j appear more than once in equation (6). In that case, we can rewrite (6) in partially bundle form,

$$\Delta x_i + \Delta x_{i+2j} = \sum_{k=1}^{q} N_k \Delta z'_k \ , \tag{9}$$

where

$$\Delta z'_k = \sum_{\{l|N_k = M_l\}} \Delta z_l \ . \tag{10}$$

In this case, we also have to consider the weight condition for XOR of multiple element to treat $\Delta z'_k = \sum_{\{l|N_k=M_l\}} \Delta z_l$. Let $j \geq 1$, W_3 be $w_h(\sum_{i=1}^{j} \Delta a_i)$, and let w_{max} be $max(w_h(\Delta a_1), w_h(\Delta a_2), .., w_h(\Delta a_j))$, we obtain the range of W_3 as,

$$max(0, 2w_{max} - \sum_{i=1}^{j} w_h(\Delta a_i)) \leq W_3 \leq min(m, \sum_{i=1}^{j} w_h(\Delta a_i)) \tag{11}$$

Then this weight condition W_3 can be used to determine the weight condition of W_2 for the equation (9).