Comparing Substitution Boxes of the Third Generation GSM and Advanced Encryption Standard Ciphers

Sedat Akleylek, Melek D. Yücel

Abstract— **Block ciphers have an important position in most of the security protocols. Security of these protocols depends on the security of the underlying block ciphers. Resistance to cryptanalytic techniques of block ciphers starts from their smallest component substitution boxes (s-boxes) which provide diffusion and confusion. Characteristics of s-boxes are the measure of strength to cryptanalytic techniques. Diffusion and confusion properties of the substitution boxes (s-boxes) of the three ciphers; i) KASUMI, the standard algorithm for the 3rd Generation Partnership Project in order to be used in the 3rd Generation mobile systems, ii) MISTY1, the predecessor of KASUMI and iii) RIJNDAEL, the Advanced Encryption Standard, are investigated. S-box characteristics of MISTY1, KASUMI and RIJNDAEL are evaluated according to strict avalanche criterion, linear approximation table and XOR table distributions. Experimental results show that there are some differences between s-boxes.**

***Index Terms*—LAT, SAC, s-box, XOR**

I. Introduction

SHANNON presented the principles of diffusion and confusion in 1949 [1]. To design a cipher according to the principle of diffusion means that one can design it to ensure that "the statistical structure of plaintext which leads to its redundancy is dissipated into long term statistics". That is every bit of the ciphertext should depend on every bit of the plaintext and every bit of the key. The higher the diffusion, the more output bits can be affected by a certain input bit. Strict avalanche criterion, the combination of completeness and avalanche criteria, is the measure of diffusion.

To design a cipher according to the principle of confusion means that one can design it so as "to make the relation between the simple statistics of ciphertext and the simple description of key a very complex and involved one". Ideally, every bit in the key influences every bit of the ciphertext so that it depends on all parts of the key, and this dependence appears to be random. Difference to linear functions measures the quality of confusion. The security of cryptographic algorithms depends upon the strength, namely the diffusion and confusion properties of the constituting Boolean functions. Boolean functions are the main building blocks of most cipher systems. A Boolean function produces a single bit result for each possible combination of values form many Boolean variables, namely a Boolean function of n variables is a function $f(x): F_2^n \rightarrow F_2$. The Boolean field consists of the values {0,1}.

An n×n substitution box (s-box) is a mapping function, $S(x): F_2^n \rightarrow F_2^n$, which maps n-bit inputs to n-bit outputs and can be viewed as consisting of n Boolean functions. Since diffusion and confusion properties are added to a block cipher by s-boxes, some of the research about the cryptographic strength of block ciphers has been focused on some properties of the s-boxes such as strict avalanche criterion, linear approximation table (LAT) and exclusive or (XOR) table distributions.

The three algorithms, whose s-boxes are compared in this work, are the block ciphers MISTY1, KASUMI and RIJNDAEL. MISTY1 [2] is an encryption algorithm developed by Mitsubishi Electric and said to be one of the most reliable and secure encryption tools. MISTY1 was recommended for Japanese government use by the Cryptography Research and Evaluation Committee (CRYPTREC project) [3] in 2003. MISTY1's security capabilities are used as the base for KASUMI [4].

Within the security architecture of the Third Generation Partnership Project (3GPP) system there are two standardized algorithms: A confidentiality algorithm f8 and an integrity algorithm f9 [5]. Each of these algorithms is based on the KASUMI block cipher, which is derived from MISTY1. KASUMI is an eight round Feistel network that produces a 64-bit output from a 64-bit input, under the control of a 128-bit key.

Both MISTY1 and KASUMI use two different s-boxes : S_7, which maps a 7-bit input to a 7-bit output and S_9, which maps a 9-bit input to a 9-bit output . The s-boxes S_7 and S_9 are obtained as linear transforms of power functions over the corresponding fields, with Kasami's exponents [6]. S_7 and S_9 were designed with function $x \rightarrow x^{81}$ in F_2^7 and

Manuscript received September 25, 2007; revised November 23, 2007.

Sedat Akleylek is with the Institute of Applied Mathematics, Middle East Technical University, Ankara, 06531 Turkey. (e-mail:akleylek@metu.edu.tr).

Melek D. Yücel is with the Electrical Engineering Department and Institute of Applied Mathematics, Middle East Technical University, Ankara, 06531 Turkey. (e-mail: melekdy@metu.edu.tr).

$x \to x^5$ in F_2^9, respectively. These s-boxes are chosen according to minimum linear/differential probability and high algebraic degree, to resist linear/differential and algebraic attacks.

The third cipher, RIJNDAEL designed by Joan Daemen and Vincent Rijmen was selected as the Advanced Encryption Standard (AES) by NIST [7] in 2000. RIJNDAEL is a block cipher that produces 128, 192 or 256-bit outputs from 128, 192 or 256-bit inputs under the control of 128, 192 or 256-bit keys.

RIJNDAEL uses S_8, which maps a 8-bit input to a 8-bit output. RIJNDAEL's s-box is designed using the function $x \to x^{-1}$ in F_2^8 [8]. The s-box of RIJNDAEL first calculates $x \to x^{-1}$ in F_2^8. The resulting inverse is transformed by an affine transformation to produce the output.

In this study, we present and compare the experimental results for the strict avalanche criterion, LAT and XOR table distributions, for the s-boxes of MISTY1, KASUMI and RIJNDAEL. In Appendix, we give the results in details with tables and figures.

II. DEFINITIONS

A. Avalanche Criterion

The idea of avalanche effect was first defined by Feistel [9]. For a given transformation to exhibit the avalanche effect, an average of one half of the output bits should change whenever a single input bit is complemented. More formally, a function $S(x): F_2^n \to F_2^n$ satisfies the avalanche criterion if whenever one input bit is changed, on the average half of the output bits change.

Let $A^{e_i} = S(x) \oplus S(x \oplus e_i) = (a_1^i, a_2^i, ..., a_n^i)$ be the avalanche vector for all $e_i \in F_2^n$ such that $wt(e_i) = 1$, where $1 \le i \le n$. Then, the avalanche criterion is satisfied when the parameter defined in [10] as

$AVAL(e_i) = \sum_{j=1}^{n} \sum_{x \in F_2^n} a_j^i$ becomes equal to $n \cdot 2^{n-1}$.

$\overline{AVAL}(e_i) = \frac{1}{n \cdot 2^n} \sum_{j=1}^{n} \sum_{x \in F_2^n} a_j^i$ is called the normalized avalanche criterion. If it is equal to 1/2 for all i, then $S(x): F_2^n \to F_2^n$ satisfies the avalanche criterion.

B. Strict Avalanche Criterion

Webster and Tavares combined the criteria of completeness and avalanche into strict avalanche criterion (SAC) [11]. If a cryptographic function $S(x): F_2^n \to F_2^n$ is to satisfy the SAC, the change of the i^{th} input bit results in the change of the j^{th} output bit exactly half of all possible input vectors, which is the probability that the j^{th} output bit is complemented is 1/2.

Let $A^i = S(x) \oplus S(x \oplus i) = (a_1^i, a_2^i, ..., a_n^i)$ be the avalanche vector, for $i \in F_2^n - \{0\}$. Then, if $SAC(i, j) = \sum_{x \in F_2^n} a_j^i$ is equal to 2^{n-1} for all i and j, then the n×n s-box or the vector Boolean function $S(x): F_2^n \to F_2^n$ satisfies the SAC.

The measure D_j^i, the normalized distance to SAC, is defined in [10] and it can be used to indicate how close the function $S(x)$ is, to satisfy SAC.

$$D_j^i = \frac{1}{2^{n-1}} \left(2^{n-1} - \sum_{x \in F_2^n} a_j^i \right) \qquad (1)$$

where i is any n bit vector, $1 \le i \le 2^n - 1$, a_j^i is the j^{th} avalanche variable, $1 \le j \le n$, of the avalanche vector A^i, and x is the input vector. If the strict avalanche criterion is satisfied, then $|D_j^i| = 0$ for all output bits. In the worst case, $\max |D_j^i| = 1$, where the maximum is computed over all i and j. It should be noted that satisfying this criterion is not sufficient to ensure the security of the cipher.

C. Linear Approximation Table (LAT) Distribution

Linear cryptanalysis [12] is a known plaintext attack that is based on effective linear approximate relations between the plaintext, ciphertext and the key.

Definition 1. Let $x, y \in F_2^n$ and $S(x): F_2^n \to F_2^n$. Each element of the Linear Approximation Table is defined as

$$\underset{a,c \in F_2^n}{LAT}(a, c) = \#\{x \mid c \bullet S(x) = a \bullet x\} - 2^{n-1}, \qquad (2)$$

where a and c are respectively the row and column indices and $\bullet$ denotes the dot product of vectors. After normalizing LAT elements, $\frac{\underset{a,c \in F_2^n}{LAT}(a,c)}{2^n}$, one can get probability bias. Probability bias helps us to determine deviation from 1/2 and varies in the interval [-1/2, 1/2]. LAT is an important tool to measure the security of s-boxes against linear cryptanalysis. Large elements of LAT are not desired since they indicate high probability of linear relations between the input and the output.

D. Exclusive Or (XOR) Table Distribution

Differential cryptanalysis [13] is a chosen plaintext attack, which uses the propagation of input differences to output differences in iterated transforms. In other words, it exploits

the high propagation probability of certain occurrences of plaintext differences to the last round input difference of the cipher.

Definition 2. Let $x, y \in F_2^n$ and $S(x) : F_2^n \rightarrow F_2^n$. Let two inputs to the system be x', x'' with the corresponding outputs y', y'' respectively. The input and output differences are given by $\Delta x = x' \oplus x''$ and $\Delta y = y' \oplus y''$, respectively. Then, the XOR table can be constructed by using

$$XOR(\Delta x, \Delta y) = \#\{x \mid S(x) \oplus S(x \oplus \Delta x) = \Delta y\} \quad (3)$$

The rows of the matrix, Δx, represent the change in the output of the s-box. The parameter $\max_{\Delta x, \Delta y \neq 0} XOR(\Delta x, \Delta y) = \delta$ is called the differential uniformity. XOR table of an s-box gives information about the security of the block cipher against differential cryptanalysis. If differential uniformity is large, this is an indication of an insecure block cipher. Differential probability can be calculated by $DP(\Delta x, \Delta y) = \frac{\delta}{2^n}$.

III. RESULTS

A. Test Results for Strict Avalanche Criterion

SAC values constitute a table of size 127×7 for the 7×7 s-boxes, 255×8 for the 8×8 s-box and 511×9 for the 9×9 s-boxes, whose elements are calculated by (1). When the s-boxes of KASUMI, MISTY1 and RIJNDAEL are compared, most of the normalized distance to SAC values is found as 0 in MISTY1 and KASUMI, while in RIJNDAEL, all values are distributed around 0. Table 1 summarizes the maximum normalized distance obtained over all i and j the corresponding input difference vector i, for the five different s-boxes used by the three block ciphers. In Table 1 only one value is given in the corresponding i column.

Table 1: Normalized Distance to SAC for the s-boxes

S-box	$\left\|D_j^i\right\|_{\max}$	Corresponding i
MISTY 7×7 s-box	0.125	127
KASUMI 7×7 s-box	0.125	127
MISTY 9×9 s-box	1	128
KASUMI 9×9 s-box	1	128
RIJNDAEL 8×8 s-box	0.125	72

Figures of normalized distance to SAC of the s-boxes are given in Appendix A. It is observed that for 7×7 s-boxes SAC gets the highest value for each bit of avalanche vector. SAC gets the highest value for at most two bits of avalanche vector for Rijndael 8×8 s-box. For 9×9 s-box of KASUMI there are only two values {0, 1}. On the other side, For MISTY1's 9×9 s-box there are three values {0, 0.5, 1}.

B. Test Results for LAT Distribution

LAT is a table of size 128×128 for the 7×7 s-boxes, 256×256 for the 8×8 s-boxes and 512×512 for the 9×9 s-boxes, whose elements are calculated by (2). When we compare 7×7 s-boxes and 9×9 s-boxes of MISTY1 and KASUMI according to LAT, it is seen that LAT values are uniformly distributed and the highest absolute value of LAT elements is 8 for 7×7 s-boxes of MISTY1 and KASUMI, and 16 for 9×9 s-boxes of MISTY1 and KASUMI. However, MISTY1's is more changeable. In RIJNDAEL 8×8 s-box, on the contrary MISTY1's and KASUMI's s-boxes absolute LAT values differs between 2 and 16. Table 2 shows the distribution percentages of LAT elements for the three s-boxes.

Table 2: Distribution Percentages of LAT Elements for the three s-boxes

S-box Size / Percentages of	7×7	8×8	9×9
0's	0.50	0.07	0.50
|2|'s	0	0.18	0
|4|'s	0	0.14	0
|6|'s	0	0.15	0
|8|'s	0.49	0.13	0
|10|'s	0	0.09	0
|12|'s	0	0.14	0
|14|'s	0	0.06	0
|16|'s	0	0.01	0.49

According to probability biases, there is no difference between 7×7 and 8×8 s-boxes. On the other hand, 9×9 s-boxes have the best probability bias for linear cryptanalysis. Table 3 gives the probability biases for the three s-boxes.

Table 3: Probability Biases for the three s-boxes

S-box Size	7×7	8×8	9×9
Probability Bias	1/16	1/16	1/32

Tables of LAT distributions of all s-boxes for single bit input and output differences are tabulated in Appendix B. It is

observed from Table B-5 that there is only one |16| in each row and column for KASUMI's 9×9 s-box.

C. Test Results for XOR Table

The XOR table is a matrix of size 128×128 for the 7×7 s-boxes, 256×256 for the 8×8 s-boxes and 512×512 for the 9×9 s-boxes, whose elements are calculated by (3). Differential uniformity of 7×7 and 9×9 s-boxes is 2, on the other hand differential uniformity of 8×8 s-box is 4. However, in RIJNDAEL 8×8 s-box XOR values differ from 0, 2 and 4. Every row and column except 1th row contains exactly one 4. It can be said that XOR values of all s-boxes are uniformly distributed. In other words, all s-boxes have the equally probable XOR table distribution. However, 8x8 s-box is less equally probable than the others. Table 4 summarizes the number of XOR table elements for the three s-boxes except 1th row.

Table 4 : XOR Table Elements for Each Row of the three s-boxes

S-box Size	7×7	8×8	9×9
Number of			
0's	2^{n-1}	$2^{n-1}+1$	2^{n-1}
2's	2^{n-1}	$2^{n-1}-2$	2^{n-1}
4's	0	1	0

Table 5 shows differential probabilities of the three s-boxes. According to differential probability measure 9×9 s-boxes give the best result.

Table 5: Differential Probability for the three s-boxes

S-box Size	7×7	8×8	9×9
Differential Probability	1/64	1/64	1/256

Tables of XOR table distributions of all s-boxes for single bit input and output differences are tabulated in Appendix C. It is observed from Table C-5 that there is only one 2 in each row and column for KASUMI's 9×9 s-box.

IV. CONCLUSION

Experimental results for the strict avalanche criterion show that s-boxes of MISTY1 7×7, KASUMI 7×7 and RIJNDAEL are much better than the s-boxes of MISTY1 9×9 and KASUMI 9×9 since they satisfy this criterion within very small deviation such as 0.125. On the other hand, most of the values in MISTY1's 9×9 and KASUMI's 9×9 s-boxes are 0, but the maximum value is 1. Figure A-3 shows that normalized distance to SAC values of RIJNDAEL's 8×8 s-box are more uniformly distributed. LAT and XOR table distributions show that KASUMI's 9×9 s-box is superior than MISTY1's 9×9 s-box according to Table B-5 and C-5. It is observed from Table B-5 and C-5, LAT and XOR table distributions have zeros at the same places for KASUMI's 9×9 s-box. This implies that there is a resemblance between LAT and XOR table distributions. It can be concluded that s-boxes of MISTY1 7×7, KASUMI 7×7 have essentially same cryptographic properties according to SAC, LAT and XOR table distribution. On the other hand, there are significant differences between MISTY's 9×9 s-box and KASUMI's 9×9 s-box although they have constructed in the same way. Although there are differences between s-boxes, these results may not affect overall security of ciphers since effectiveness of Linear and Differential cryptanalysis depends upon probability biases and differential probabilities which are very small to attack for all s-boxes.

APPENDIX A

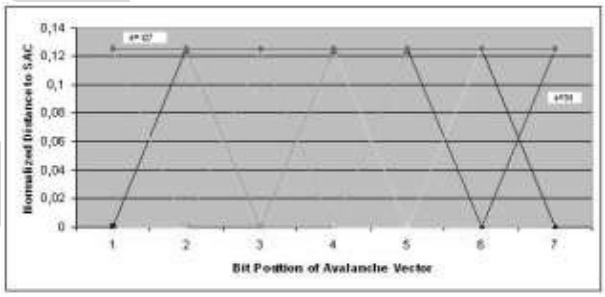

Fig. A-1 : Normalized Distance to SAC of MISTY1's 7×7 s-box

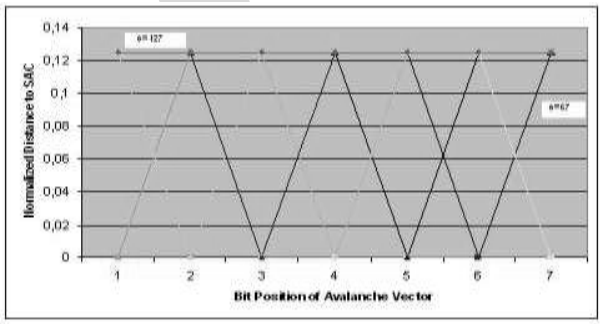

Fig. A-2 : Normalized Distance to SAC of KASUMI's 7×7 s-box

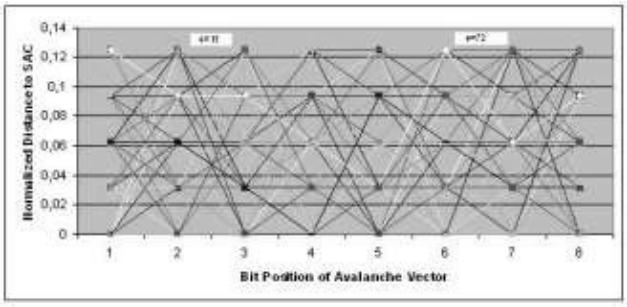

Fig. A-3 : Normalized Distance to SAC of RIJNDAEL's 8×8 s-box

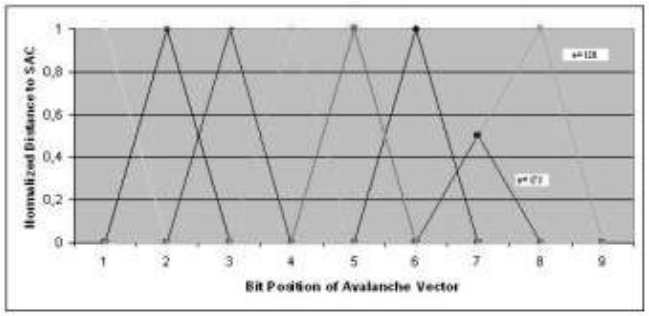

Fig. A-4 : Normalized Distance to SAC of MISTY1's 9×9 s-box

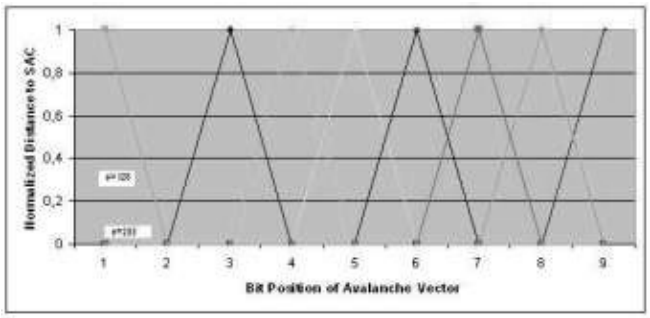

Fig. A-5 : Normalized Distance to SAC of KASUMI's 9×9 s-box

APPENDIX B

Table B-1: LAT Distribution of MISTY1's 7×7 s-box for Single bit Input and Output Differences

Output Sum / Input Sum	1	2	4	8	16	32	64
1	-8	0	8	0	0	0	0
2	0	0	0	-8	-8	0	8
4	0	-8	8	0	0	8	8
8	0	-8	0	8	0	8	8
16	-8	0	8	0	0	-8	0
32	0	0	0	-8	-8	0	0
64	0	-8	0	0	0	8	8

Table B-2: LAT Distribution of KASUMI's 7×7 s-box for Single bit Input and Output Differences

Output Sum / Input Sum	1	2	4	8	16	32	64
1	8	0	-8	0	-8	0	0
2	0	-8	0	8	0	0	0
4	-8	0	0	0	0	-8	8
8	8	8	-8	0	-8	0	0
16	8	0	-8	0	-8	-8	0
32	0	-8	0	8	-8	0	0
64	0	0	0	0	0	-8	8

Table B-3: LAT Distribution of RIJNDAEL's 8×8 s-box for Single bit Input and Output Differences

Output Sum / Input Sum	1	2	4	8	16	32	64	128
1	12	0	14	12	8	-4	4	12
2	2	8	2	6	-2	8	-16	-2
4	-8	2	6	6	12	-16	2	-2
8	2	2	4	0	12	6	2	4
16	-12	-2	-6	-2	-8	-10	0	8
32	6	-10	-2	-12	2	0	-8	12
64	4	-4	-12	16	6	-8	-12	4
128	-12	-12	16	14	-8	-12	-4	-4

Table B-4: LAT Distribution of MISTY1's 9×9 s-box for Single bit Input and Output Differences

Output Sum / Input Sum	1	2	4	8	16	32	64	128	256
1	-16	0	0	16	0	0	0	0	0
2	-16	0	0	0	16	0	0	0	0
4	-16	0	0	0	0	16	0	0	0
8	-16	0	0	0	0	0	-16	0	0
16	-16	0	0	0	0	0	0	0	-16
32	-16	0	0	0	0	0	0	-16	0
64	-16	0	0	0	0	0	0	0	0
128	-16	-16	0	0	0	0	0	0	0
256	-16	0	16	0	0	0	0	0	0

Table B-5: LAT Distribution of KASUMI's 9×9 s-box for Single bit Input and Output Differences

Output Sum / Input Sum	1	2	4	8	16	32	64	128	256
1	0	0	0	0	0	0	16	0	0
2	0	0	16	0	0	0	0	0	0
4	0	0	0	0	0	-16	0	0	0
8	-16	0	0	0	0	0	0	0	0
16	0	0	0	0	16	0	0	0	0
32	0	0	0	16	0	0	0	0	0
64	0	-16	0	0	0	0	0	0	0
128	0	0	0	0	0	0	0	0	16
256	0	0	0	0	0	0	0	-16	0

APPENDIX C

Table C-1: XOR Table Distribution of MISTY1's 7×7 s-box for Single bit Input and Output Differences

Δy / Δx	1	2	4	8	16	32	64
1	2	0	0	0	0	0	0
2	0	0	0	2	0	0	0
4	0	0	0	0	0	2	0
8	0	0	0	0	0	2	2
16	0	0	2	0	0	2	2
32	0	0	2	0	2	2	2
64	0	2	2	0	2	2	2

Table C-2: XOR Table Distribution of KASUMI's 7×7 s-box for Single bit Input and Output Differences

Δy / Δx	1	2	4	8	16	32	64
1	2	0	2	2	2	2	0
2	2	0	0	2	2	2	0
4	2	0	0	0	2	2	0
8	2	0	0	0	2	0	0
16	2	0	0	0	0	0	0
32	0	2	0	0	0	0	0
64	0	0	0	0	0	0	2

Table C-3: XOR Table Distribution of RIJNDAEL's 8×8 s-box for Single bit Input and Output Differences

Δy / Δx	1	2	4	8	16	32	64	128
1	0	0	0	0	0	0	0	0
2	2	0	2	2	0	0	0	2
4	0	2	2	2	0	0	2	0
8	2	0	0	2	2	0	0	2
16	0	2	2	2	0	0	0	2
32	0	2	0	2	0	0	2	0
64	0	0	2	0	2	0	2	2
128	0	0	0	2	2	2	2	2

Table C-4: XOR Table Distribution of MISTY1's 9×9 s-box for Single bit Input and Output Differences

Δy / Δx	1	2	4	8	16	32	64	128	256
1	0	0	0	2	0	0	0	0	0
2	0	0	0	0	2	0	0	0	0
4	0	0	0	0	0	2	0	0	0
8	0	0	0	0	0	0	2	0	0
16	0	0	0	0	0	0	0	0	2
32	0	0	0	0	0	0	0	2	0
64	0	2	2	2	2	2	2	2	2
128	0	2	0	0	0	0	0	0	0
256	0	0	2	0	0	0	0	0	0

Table C-5: XOR Table Distribution of KASUMI' 9×9 s-box for Single bit Input and Output Differences

Δy / Δx	1	2	4	8	16	32	64	128	256
1	0	0	0	0	0	0	2	0	0
2	0	0	2	0	0	0	0	0	0
4	0	0	0	0	0	2	0	0	0
8	2	0	0	0	0	0	0	0	0
16	0	0	0	0	2	0	0	0	0
32	0	0	0	2	0	0	0	0	0
64	0	2	0	0	0	0	0	0	0
128	0	0	0	0	0	0	0	0	2
256	0	0	0	0	0	0	0	2	0

REFERENCES

[1] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell System Technical Journal*, vol.28-4, pp. 656—715, 1949.

[2] M. Matsui, "New Block Encryption MISTY," *FSE 4th International Workshop, Lecture Notes in Computer Science*, vol.1267, pp. 54-68, *Springer Verlag*, 1997.

[3] CRYPTREC, Cryptography Research and Evaluation Committee. Avaliable : http://www.ipa.go.jp/security/enc/CRYPTREC/index-e.html

[4] 3rd Generation Partnership Project, "Technical Specification Group Services and System Aspects, 3G Security, Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 2: KASUMI Specification, V.3.1.1," 2001.

[5] ETSI/SAGE, "KASUMI Specification, Part of the Specification of the 3GPP Confidentiality and Integrity Algorithms," 1999.

[6] H. Dobbertin, "Almost Perfect nonlinear Power Functions on GF(2n): The Welch case," *IEEE Transactions on Information Theory*, Vol. 45, NO.4, 1999.

[7] J. Daemen and V. Rijmen, "AES Proposal : Rijndael," *NIST Publication*, 1999.

[8] K. Nyberg, "Differentially uniform mappings for cryptography," *EUROCRYPT'93, Lecture Notes in Computer Science*, vol. 765, *Springer Verlag*, pp. 55–64, 1994.

[9] H. Feistel, "Cryptography and Computer Privacy," *Scientific American*, Volume 228, No:5, pp.15-23, 1973.

[10] E. Aras, "Analysis of Security Criteria for Block Ciphers," M.S. Thesis, Middle East Technical University, Turkey, 1999.

[11] A. Webster and S. Tavares, "On the Design of s-boxes," Advances in Cryptology, Proc. *EUROCRYPT'85, Springer Verlag*, pp.523-534, 1986.

[12] M. Matsui, "Linear Cryptanalysis Method for DES cipher," *EUROCRYP' 93, Lectures Notes in Computer Science* no. 765, *Springer Verlag*, pp. 386-397, 1994.

[13] E. Biham and A. Shamir, "Differential Cryptanalysis of DES-Like Cryptosystems," *Journal of Cryptology*, volume:4 pp. 3-72, 1991.