Linear Approximation Analysis: An improved Technique for Linear Cryptanalysis of 4-bit Bijective Crypto S-Boxes.

Sankhanil Dey[1], Ranjan Ghosh[2].
sdrpe_rs@caluniv.ac.in[1], rghosh47@gmail.com[2].
Institute of Radio Physics and Electronics,
University of Calcutta.

Abstract. 4-bit Linear Relations play an important role in Cryptanalysis of 4-bit Bijective S-Boxes. Count of existence of all 4-bit Linear Relations, for all of 16 input and output 4-bit bit patterns of 4-bit Bijective S-Boxes said as S-Boxes has been reported in Linear Cryptanalysis of 4-bit S-Boxes. In this paper a brief review of this cryptanalytic method for 4-bit S-Boxes has been introduced in a very lucid and conceptual manner. A new Analysis to search for the existing Linear Approximations among the input Boolean Functions (BFs) and output BFs of a particular 4-bit S-Box has also been introduced in this paper. The search is limited to find the existing Linear Relations or Approximations in the contrary to count the number existence among all 16 4-bit input and output bit patterns for all possible linear approximations.

1. **Introduction.** The Exclusive-Or or Xor operation is defined to be linear operation in cryptography. Linear operations are used to give two exact values, 0 and 1, in operation between two same and different bits respectively in Boolean Logic or Switching Logic. So if a linear relation exists between all 4-bit plain text bit pattern and the corresponding cipher-text 4-bit bit pattern then the existing relation between them is easy to determine. The idea of using linear relations to analyze the randomization property of a cipher was introduced by Matsui in 1994 for cryptanalysis reduced round DES cipher [1]. Later Heys [2] extended the idea towards 4-bit S-Boxes in his tutorial on linear and Differential Cryptanalysis of 4-bit S-Boxes.

A 4-bit S-Box consists of 16 array elements whose indices are considered as 4-bit inputs corresponding to sequential hex values from of 0 to f. The output data corresponding to each array indices are supposed to have 4-bit sequential or non-sequential hex values between 0 and f. Such an S-Box with 4-bit input and 4-bit output are called a Bijective S-Box [3]. Non-bijective 4-bit S-Boxes are those whose inputs may consist of number of bits more than four bits. For all 4-bit Bijective S-Boxes, the four input vectors are same and the output would be composed of four Boolean functions (BFs) giving four 16-bit output column vectors whose row-wise 4 bits assume hex values lying between 0 and f. The number of possible S-boxes is obtained as factorial 16 (16!) following the permutation of 16 hex digits between 0 and f. a 4-bit S-box can be represented by a 4-valued Boolean function following the norms of presentation of multi-valued Boolean function [4].

For Linear Cryptanalysis of 4-bit S-Boxes, every 4-bit linear relations are tested for a particular 4-bit S-Box. The presence of each 4-bit unique linear relation is checked by satisfaction of each of them for all 16, 4-bit unique input bit patterns and corresponding 4-bit output bit patterns, generated from the index of each element and each element respectively of that particular 4-bit S-Box. If they are satisfied 8 times out of 16 operations for all 4-bit unique input bit patterns and corresponding 4-bit output bit patterns, then the existence of the 4-bit linear equation is at a stake, since the probability of presence and absence of a 4-bit linear relation both are (= 8/16) ½. If a 4-bit linear equation is satisfied 0 times then it can be concluded that the given 4-bit linear relation is absent for that particular 4-bit bijective S-Box. If a 4-bit linear equation is satisfied 16 times then it can also be concluded that the given 4-bit linear relation is present for that particular 4-bit S-Box. In both the cases full information is adverted to the cryptanalysts. The concept of Probability Bias was introduced to predict the randomization ability of that 4-bit S-Box from the probability of presence or absence of unique 4-bit linear relations. The result is better for cryptanalysts if the probability of presence or absences of unique 4-bit linear equations are far away from ½ or near to 0 or 1. If the probabilities of presence or absence, of all unique 4-bit linear relations are ½ or close to ½, then the 4-bit S-Box is said to be linear cryptanalysis immune, since the existence of maximum 4-bit linear relations for that 4-bit S-Box is hard to predict. Heys also introduced the concept of Linear Approximation Table (LAT) in which the numbers of times, each 4-bit unique linear relation have been satisfied for all 16, unique 4-bit input bit patterns and corresponding 4-bit output bit patterns, of a crypto S-box are noted. The result is better for a cryptanalysts if the numbers of 8s in the table are less. If numbers of 8s are much more than the other numbers in the table then the 4-bit S-Box is said to be more linear cryptanalysis immune.

In this paper a new technique to find the existing Linear Relations or Linear Approximations for a particular 4-bit Bijective S-Box has been introduced. If the nonlinear part of the ANF equation of a 4-bit output BF is absent or calculated to be 0 then the equation is termed as a Linear Relation or Approximation. Searching for number of existing linear relations through this method is ended up with Number of Existing Linear Relations. I.e. the goal to conclude the security of a 4-bit bijective S-Box is attended in a very lucid manner by this method.

The review of and Algebraic Normal Form of 4-bit BFs and Linear Cryptanalysis of 4-bit Bijective S-Boxes have been given or introduced in a very lucid manner in section 2. The new Cryptanalysis method or Linear Approximation Analysis has been described in section 3. The algorithm is given in section 4. The analysis of results and security Criterion for 4-bit Bijective S-Boxes has been given in section 5. The conclusion has been made in section 6. The references are elaborated in section 7. The analysis of 32 4-bit Bijective Crypto S-Boxes of Data Encryption Standard has been shown in Appendix.

2. Review of Algebraic Normal Form and Linear Cryptanalysis.

The review of algebraic Normal form or ANF of 4-bit BFs is given in subsection 2.1. and a lucid review of Linear Cryptanalysis of 4-bit Bijective Crypto S-Boxes is given in subsection 2.2.

2.1 A review of Boolean Functions (BF) and its Algebraic Normal Form (ANF)

A 4-bit Boolean Function (BF) accepts 4 bits as input $\{x_1x_2x_3x_4\}$having 16 combinations of decimal values varying between 0 and 15 and provides 1-bit output for each combination of input. The input-output relation is given in a Truth Table which provides 16-bit output vector corresponding to four 16-bit input vectors, each one attached to x_1, x_2, x_3 and x_4. The 4-bit BF is a mapping from $(0,1)^4$ to $(0,1)^1$and its functional relation, F(x) can be expressed in Algebraic Normal Form (ANF) with 16 coefficients as given in eq. (1) below,

$$F(x) = a_0 + (a_1.x_1 + a_2.x_2 + a_3.x_3 + a_4.x_4) + (a_5.x_1.x_2 + a_6.x_1.x_3 + a_7.x_1.x_4 + a_8.x_2.x_3 + a_9.x_2.x_4 + a_{10}.x_3.x_4) + + (a_{11}.x_1.x_2.x_3 + a_{12}.x_1.x_2.x_4 + a_{13}.x_1.x_3.x_4 + a_{14}.x_2.x_3.x_4) + a_{15}.x_1.x_2.x_3.x_4 \quad \ldots \quad \ldots \quad (1)$$

where x represents the decimal value or the hex value of 4 input bits represented by $\{x_1x_2x_3x_4\}$, BF assumes 1-bit output, '**.**' and '+' represent AND and XOR operations respectively. Here a_0 is a constant coefficient, (a_1 to a_4) are 4 linear coefficients, and (a_5 to a_{15}) are 11 nonlinear coefficients of which (a_5 to a_{10}) are 6 non-linear coefficients of 6 terms with 2-AND-operated-input-bits, (a_{11} to a_{14}) are 4 nonlinear coefficients of 4 terms with 3-AND-operated-input-bits and a_{15} is a non-linear coefficient of one term with 4-AND-operated-input-bits. The 16 binary ANF coefficients, from a_0 to a_{15} are marked respectively as anf.bit0 to anf.bit15 in ANF representation and are evaluated from the 16-bit output vector of a BF designated as bf.bit0 to bf.bit15 using the following relations as given in eq.(2),

anf.bit0 = bf.bit0;
anf.bit1 = anf.bit0 + bf.bit8;
anf.bit2 = anf.bit0 + bf.bit4;
anf.bit3 = anf.bit0 + bf.bit2;
anf.bit4 = anf.bit0 + bf.bit1;
anf.bit5 = anf.bit0 + anf.bit1 + anf.bit2 + bf.bit12;
anf.bit6 = anf.bit0 + anf.bit1 + anf.bit3 + bf.bit10;
anf.bit7 = anf.bit0 + anf.bit1 + anf.bit4 + bf.bit9;
anf.bit8 = anf.bit0 + anf.bit2 + anf.bit3 + bf.bit6;
anf.bit9 = anf.bit0 + anf.bit2 + anf.bit4 + bf.bit5;
anf.bit10 = anf.bit0 + anf.bit3 + anf.bit4 + bf.bit3;
anf.bit11 = anf.bit0 + anf.bit1 + anf.bit2 + anf.bit3 + anf.bit5 + anf.bit6 + anf.bit8 + bf.bit14;
anf.bit12 = anf.bit0 + anf.bit1 + anf.bit2 + anf.bit4 + anf.bit5 + anf.bit7 + anf.bit9 + bf.bit13;
anf.bit13 = anf.bit0 + anf.bit1 + anf.bit3 + anf.bit4 + anf.bit6 + anf.bit7 + anf.bit10 + bf.bit11;
anf.bit14 = anf.bit0 + anf.bit2 + anf.bit3 + anf.bit4 + anf.bit8 + anf.bit9 + anf.bit10 + bf.bit7;
anf.bit15 = anf.bit0 + anf.bit1 + anf.bit2 + anf.bit3 + anf.bit4 + anf.bit5 + anf.bit6 + anf.bit7
+ anf.bit8 + anf.bit9 + anf.bit10 + anf.bit11 + anf.bit12 + anf.bit13 + anf.bit14 + bf.bit15 … (2)

The DEBF (Decimal Equivalent of BF) varies from 0 through 65535 and each decimal value is converted to a 16-bit binary output of the Boolean function from bf.bit0 through bf.bit15. Based on the binary output of a BF, the ANF coefficients from anf.bit0 through anf.bit15 are calculated sequentially using eq. (2).

2.2 A Review on Linear Cryptanalysis of 4-bit Bijective Crypto S-Box [2].

The given 4-bit Bijective crypto S-Box has been described in sub-section 2.2.1. The relation Between 4-bit S-Boxes and 4 bit BFs, The Linear Approximations are described in sub-section 2.2.2 and 2.2.3 respectively. LAT or Linear Approximation Table has been illustrated in sec 2.2.4.

2.2.1. 4-bit Crypto S-Boxes: A 4-bit bijective Crypto S-Box can be written as Follows, where the each element of the first row of Table.1, entitled as index, are the position of each element of the S-Box within the given S-Box and the elements of the 2nd row, entitled as S-Box, are the elements of the given Substitution Box. It can be concluded that the 1st row is fixed for all possible bijective crypto S-Boxes. The values of each element of the 1st row are distinct, unique and vary between 0 and F. The values of the each element of the 2nd row of a bijective crypto S-Box are also distinct and unique and also vary between 0 and F. The values of the elements of the fixed 1st row are sequential and monotonically increasing where for the 2nd row they can be sequential or partly sequential or non- sequential. Here the given Substitution Box is the 1st 4-bit S-Box of the 1st S-Box out of 8 of Data Encryption Standard [5][6][7].

Row	Column	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G
1	Index	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
2	S-Box	E	4	D	1	2	F	B	8	3	A	6	C	5	9	0	7

Table.1. 4-bit bijective Crypto S-Box.

2.2.2. Relation between 4-bit S-Boxes and 4-bit Boolean Functions (4-bit BFs). Index of Each element of a 4-bit bijective crypto S-Box and the element itself is a hexadecimal number and that can be converted into a 4-bit bit sequence. From row 2 through 5 and row 7 through A of each column from 1 through G of Table.2. shows the 4-bit bit sequences of the corresponding hexadecimal numbers of the index of each element of the given S-Box and each element of the S-Box itself. Each row from 2 through 5 and 7 through A from column 1 through G constitutes a 16 bit, bit sequence that is a 4-bit BF. column 1 through G of Row 2 is termed as 4th Input BF, Row 3 is termed as 3rd Input BF, Row 4 is termed as 2nd Input BF and Row 5 is termed as 1st Input BF whereas column 1 through G of Row 7 is termed as 4th Output BF, Row 8 is termed as 3rd Output BF, Row 9 is termed as 2nd Output BF and Row A is termed as 1st Output BF [8]. The decimal equivalent of each input BF and output BF are noted at column H of respective rows.

Row	**Column**	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	**H.** Decimal
1	**Index**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**	Equivalent
2	**IBF4**	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	00255
3	**IBF3**	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	03855
4	**IBF2**	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	13107
5	**IBF1**	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	21845
6	**S-Box**	E	4	D	1	2	F	B	8	3	A	6	C	5	9	0	7	
7	**OBF4**	1	0	1	0	0	1	1	1	0	1	0	1	0	1	0	0	42836
8	**OBF3**	1	1	1	0	0	1	0	0	0	0	1	1	1	0	0	1	58425
9	**OBF2**	1	0	0	0	1	1	1	0	1	1	1	0	0	0	0	1	36577
A	**OBF1**	0	0	1	1	0	1	1	0	1	0	0	0	1	1	0	1	13965

2.2.3. 4-bit Linear Relations. The elements of Input S-box are shown under column heading 'I' and the Input Vectors are shown under field IPVs (Input Vectors) and subsequently under column headings 1, 2, 3 and 4. The 4th input vector is depicted under column heading '4', 3rd input vector is depicted under column heading '3', 2nd input vector is depicted under column heading '2' and 1st input vector is depicted under column heading '1'. The elements of Output S-box are shown under column heading 'SB' and the Output Vectors are shown under field OPBFs (Output Boolean Functions) and subsequently under column headings 1, 2, 3 and 4. The 4th Output BF is depicted under column heading '4', 3rd Output BF is depicted under column heading '3', 2nd Output BF is depicted under column heading '2' and 1st Output BF is depicted under column heading '1'.

I	**IPVs**				**S**	**OPBFs**			
	4	**3**	**2**	**1**	**B**	**4**	**3**	**2**	**1**
0	0	0	0	0	**E**	1	1	1	0
1	0	0	0	1	**4**	0	1	0	0
2	0	0	1	0	**D**	1	1	0	1
3	0	0	1	1	**1**	0	0	0	1
4	0	1	0	0	**5**	0	1	0	1
5	0	1	0	1	**9**	1	0	0	1
6	0	1	1	0	**0**	0	0	0	0
7	0	1	1	1	**7**	0	1	1	1
8	1	0	0	0	**2**	0	0	1	0
9	1	0	0	1	**F**	1	1	1	1
A	1	0	1	0	**B**	1	0	1	1
B	1	0	1	1	**8**	1	0	0	0
C	1	1	0	0	**3**	0	0	1	1
D	1	1	0	1	**A**	1	0	1	0
E	1	1	1	0	**6**	0	1	1	0
F	1	1	1	1	**C**	1	1	0	0

Table. 2. IPVs and OPVs for given S-Box

The IPEs or Input Equations are possible xored terms that can be formed using four IPVs 4, 3, 2 and 1. On the other hand OPEs are possible xored terms that can be formed using four OPVs 4, 3, 2 and 1. All IPEs and OPEs are listed under the column and also row heading (IPE = OPE) from row 2 through H and column 1 through G respectively. Each cell is a linear equation equating IPE to OPE. Such as $L_{1+2+4,2+3}$ is the linear equation formed by IPE '1+2+3' i.e. the xored combination of three IPVs 1, 2 and 3 and OPE '2+3' i.e. the xored combination of two OPBFs 2 and 3. The Example of 256 possible 4-bit Linear Equations are shown in Table 4.

Rows	Columns	1	2	3	4	5
1	**IPE = OPE**	**0**	**1**	**2**	**3**	**4**
2	**0**	$L_{0,0}$	$L_{0,1}$	$L_{0,2}$	$L_{0,3}$	$L_{0,4}$
3	**1**	$L_{1,0}$	$L_{1,1}$	$L_{1,2}$	$L_{1,3}$	$L_{1,4}$
4	**2**	$L_{2,0}$	$L_{2,1}$	$L_{2,2}$	$L_{2,3}$	$L_{2,4}$
5	**3**	$L_{3,0}$	$L_{3,1}$	$L_{3,2}$	$L_{3,3}$	$L_{3,4}$
6	**4**	$L_{4,0}$	$L_{4,1}$	$L_{4,2}$	$L_{4,3}$	$L_{4,4}$
7	**1+2**	$L_{1+2,0}$	$L_{1+2,1}$	$L_{1+2,2}$	$L_{1+2,3}$	$L_{1+2,4}$
8	**1+3**	$L_{1+3,0}$	$L_{1+3,1}$	$L_{1+3,2}$	$L_{1+3,3}$	$L_{1+3,4}$
9	**1+4**	$L_{1+4,0}$	$L_{1+4,1}$	$L_{1+4,2}$	$L_{1+4,3}$	$L_{1+4,4}$
A	**2+3**	$L_{2+3,0}$	$L_{2+3,1}$	$L_{2+3,2}$	$L_{2+3,3}$	$L_{2+3,4}$
B	**2+4**	$L_{2+4,0}$	$L_{2+4,1}$	$L_{2+4,2}$	$L_{2+4,3}$	$L_{2+4,4}$
C	**3+4**	$L_{3+4,0}$	$L_{3+4,1}$	$L_{3+4,2}$	$L_{3+4,3}$	$L_{3+4,4}$
D	**1+2+3**	$L_{1+2+3,0}$	$L_{1+2+3,1}$	$L_{1+2+3,2}$	$L_{1+2+3,3}$	$L_{1+2+3,4}$
E	**1+2+4**	$L_{1+2+4,0}$	$L_{1+2+4,1}$	$L_{1+2+4,2}$	$L_{1+2+4,3}$	$L_{1+2+4,4}$
F	**1+3+4**	$L_{1+3+4,0}$	$L_{1+3+4,1}$	$L_{1+3+4,2}$	$L_{1+3+4,3}$	$L_{1+3+4,4}$
G	**2+3+4**	$L_{2+3+4,0}$	$L_{2+3+4,1}$	$L_{2+3+4,2}$	$L_{2+3+4,3}$	$L_{2+3+4,4}$
H	**1+2+3+4**	$L_{1+2+3+4,0}$	$L_{1+2+3+4,1}$	$L_{1+2+3+4,2}$	$L_{1+2+3+4,3}$	$L_{1+2+3+4,4}$

Table.3. 256, 4-bit Linear Equations with input Equations (IPE) and output Equations (OPE).

2.2.4 Linear Approximation Table (LAT) [6].

According to Heys each linear equation is tested for each of 16 4-bit patterns shown in each row under the field IPVs and subsequently under the column headings 1, 2, 3 and 4 and the corresponding 16 4-bit patterns under field OPBFs and subsequently under the column headings 1, 2, 3 and 4. If a linear equation satisfies 8 times out of 16 then the existence of the linear equation is highly unpredictable. That is the probability is ½. If the numbers of satisfaction of each linear equation is noted in respective cells of Table.4. then it is called as Linear Approximation Table or LAT.

3. Linear Approximation Analysis:

A Bijective Crypto S-Box (1st 4-bit S-Box out of 32 4-bit S-Boxes of DES) has been described in sub-section 2.2.1. The Table for four input vectors, Output BFs and corresponding ANFs has been depicted in sub-section 3.2. The analysis has been described in sub-section 3.3. The result of Analysis has been given in sub-section 3.4.

3.2 Input Vectors (IPVs)-Output BFs (OPBFs)-Algebraic Normal Forms (ANFs). The elements of Input S-box are shown under column heading 'ISB' and the Input Vectors are shown under field IPVs (Input Vectors) and subsequently under column headings 1, 2, 3 and 4. The 4th input vector is depicted under column heading '4', 3rd input vector is depicted under column heading '3', 2nd input vector is depicted under column heading '2' and 1st input vector is depicted under column heading '1'. The elements of Output S-box are shown under column heading 'OSB' and the Output Vectors are shown under field OPBFs (Output Boolean Functions) and subsequently under column headings 1, 2, 3 and 4. The 4th Output BF is depicted under column heading '4', 3rd Output BF is depicted under column heading '3', 2nd Output BF is depicted under column heading '2' and 1st Output BF is depicted under column heading '1'. The corresponding ANFs for 4 OPBFs, OPBF-4th , OPBF-3rd, OPBF-2nd , OPBF-1st, are depicted under field 'ANFs' subsequently under heading 4, 3, 2 and 1 respectively.

ISB	IPVs	**OSB**	OPBFs	ANFs
	4321		**4321**	**4321**
0	0000	E	1110	1110
1	0001	4	0100	1010
2	0010	D	1101	0011
3	0011	1	0001	1100
4	0100	2	0010	1101
5	0101	F	1111	0110
6	0110	B	1011	0111
7	0111	8	1000	0011
8	1000	3	0011	1010
9	1001	A	1010	0110
A	1010	6	0110	1010

B	1011	C	1100	1000
C	1100	5	0101	0101
D	1101	9	1001	0010
E	1110	0	0000	1010
F	1111	7	0111	0000

Table.4. Input and Output Boolean Functions
With Corresponding ANF Coefficients of the given S-Box.

3.3 Linear Approximation Analysis (LAA). An algebraic Normal Form or ANF equation is termed as Linear Equation or Linear Approximation if the Nonlinear Part or NP (i.e. The xored value of all product terms of equation 2 for corresponding 4 bit values of IPVs, with column heading 4, 3, 2, 1) is 0 and The Linear part or LP for corresponding 4 bit values of IPVs, with column heading 4, 3, 2, 1 is equal to corresponding BF bit values. The corresponding ANF coefficients of output BFs F(4), F(3), F(2), and F(1) are given under row heading ANF(F4), ANF(F3), ANF(F2) and ANF(F1) respectively from row 2 through 5 and column 4 through J. In which Column 4 of row 2 through 5 gives the value of Constant Coefficient (a_0 according to eqn.1.) of ANF(F4), ANF(F3), ANF(F2) and ANF(F1) respectively. Column 5 through 8 of row 2 through 5 gives the value of respective Linear Coefficients more specifically a_1, a_2, a_3, a_4 (according to eqn. 1.) of ANF(F4), ANF(F3), ANF(F2) and ANF(F1). They together termed as LP or Linear Part of the respective ANF coefficients. Column 9 through J of row 2 through 5 gives the value of respective Non-Linear Coefficients more specifically a_5 to a_{15} (according to eqn. 1.) of ANF(F4), ANF(F3), ANF(F2) and ANF(F1). They together termed as NP or Non-Linear Part of the respective ANF coefficients.

The 4^{th}, 3^{rd}, 2^{nd}, 1^{st} IPV for The given S-Box are noted in the Field 'IPVs' under column heading 4, 3, 2, 1 respectively from row 8 through M of Table.8. The 4 output BFs F4, F3, F2, F1 are noted at column 4, 8, C, G from row 8 through M respectively. The corresponding LP, NP, Satisfaction (SF) values are noted at column 5 through 7, 9 through B, C through F and H to J from row 8 through M respectively.

R|C	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H	I	J
1	**Co-Effs**			**C**	**LP**				**NP**										
2	**ANF(F4)**			1	1	0	1	1	0	0	0	1	0	1	1	0	0	1	0
3	**ANF(F3)**			1	0	0	1	1	1	1	0	0	1	0	0	1	0	0	0
4	**ANF(F2)**			1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	0
5	**ANF(F1)**			0	0	1	0	1	0	1	1	0	0	0	0	1	0	0	0
6	**ID**	**IPVs**	**SB**	**F**	**LP**	**NP**	**SF**	**F**	**LP**	**NP**	**SF**	**F**	**LP**	**NP**	**SF**	**F**	**LP**	**NP**	**SF**
7		**4321**		**4**				**3**				**2**				**1**			
8	0	0000	E	1	0	0	1	1	0	0	1	1	1	0	0	0	1	0	1
9	1	0001	4	0	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0
A	2	0010	D	1	1	0	0	1	1	0	0	0	1	0	1	1	0	0	1
B	3	0011	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	0
C	4	0100	2	0	1	0	1	0	1	0	1	1	0	0	1	0	1	0	1
D	5	0101	F	1	0	0	1	1	0	1	1	1	0	1	1	1	1	0	0
E	6	0110	B	1	1	1	1	0	1	0	1	1	0	1	1	1	1	0	0
F	7	0111	8	1	0	1	1	0	0	1	1	0	0	0	0	0	1	0	1
G	8	1000	3	0	0	0	0	0	1	0	1	1	0	0	1	1	0	0	1
H	9	1001	A	1	1	0	0	0	0	0	0	1	0	1	1	0	0	1	1
I	A	1010	6	0	0	0	0	1	1	1	1	1	0	1	1	0	0	1	1
J	B	1011	C	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0
K	C	1100	5	0	0	0	0	1	1	1	1	0	1	1	1	1	1	0	0
L	D	1101	9	1	1	0	0	0	0	1	1	0	1	1	1	1	1	0	0
M	E	1110	0	0	0	0	0	0	1	0	1	0	1	1	1	0	1	1	1
N	F	1111	7	0	1	0	1	1	0	0	1	1	1	0	0	1	1	1	1

Table. 5 Linear Approximation Analysis.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3249v1 | CC BY 4.0 Open Access | rec: 13 Sep 2017, publ: 13 Sep 2017

3.4. Result

No. of LA with BF1	No. of LA with BF2	No. of LA with BF3	No. of LA with BF4
8	8	2	7

Total Number of Existing Linear Approximations: 25**.**

4. Algorithm: The Nonlinear Part for the given analysis has been termed as NP. The ANF coefficients are illustrated through array anf[16]. IPVs are termed as x_1, x_2, x_3, x_4 for IPV 1, IPV 2, IPV 3, IPV 4 respectively. The algorithm of the above analysis is given below,

Start.

Step 1. NP = (anf[5].&x_1&x_2)^(anf[6]&x_1 &x_3)+(anf[7]&x_1 &x_4)+(anf[8] &x_2 &x_3)+(anf[9]&x_2 &x_4)+(anf[10]&x_3 &x_4)(anf[11]&x_1 &x_2 &x_3)+(anf[12]&x_1 &x_2 &x_4)+(anf[13]&x_1 &x_3 &x_4) +(anf[14] &x_2 &x_3 &x_4)+(anf[15]&x_1 &x_2 &x_3 &x_4))

Step 2. LP= anf[0] ^(anf[1].&x_1)^ (anf[2].&x_1)^ (anf[3].&x_1)^ (anf[4].&x_1).

Step 3. if(NP==0&& BF($x_1x_2x_3x_4$) == LP) then Linear equation.
else Nonlinear equation.

Stop.

5. Analysis of Result and Security Criterion For 4-bit Bijective Crypto S-Boxes.

In this section The Analysis of result is described in sub-section 5.1. and The Security Criterion For 4-bit Bijective Crypto S-Boxes are described in subsection 5.2.

5.1 Analysis of Result.

The value of ${}^{n}C_r$ is maximum when the value of r is ½ of the value of n (when n is even). Here the maximum number of linear approximations is 64. So if the total satisfaction of linear equation is 32 out of 64 then the number of possible sets of 32 linear equations is largest. Means if the total satisfaction is 32 out of 64 then the number of possible sets of 32 possible linear equations is ${}^{64}C_{32}$. That is maximum number of possible sets of linear equations. If the value of total No of Linear Approximations with BF1 is closed to 32 then it is more cryptanalysis immune. Since the number of possible sets of linear equations are too large to calculate. As the value of goes close to 0 or 64 it reduces the sets of possible linear equations to search for which reduces the effort to search for the linear equations present in a particular 4-bit S-Box. In this example total satisfaction is 18 out of 64 . Which means the given 4-bit S-Box is not a good 4 bit S-Box or not a good Crypt analytically immune S-Box.

5.2. Security Criterion For 4-bit Bijective Crypto S-Boxes.

If the values of total number of Existing Linear equations for a 4-bit S-Box are 24 to 32, then the lowest numbers of sets of linear equations are 250649105469666120. This is a very large number to investigate. So the 4-bit S-Box is declared as a good 4-bit S-Box or 4-bit S-Box with good security. If it is between 16 through 23 then the lowest numbers of sets of linear equations are 488526937079580. This not a small number to investigate in today's computing scenario so the S-boxes are declared as medium S-Box or S-Box with medium security. The 4-bit S-Boxes having existing linear equations less than 16 are declared as Poor 4-bit S-Box or vulnerable to cryptanalytic attack.

6. Conclusion.

From this analysis it concluded that the algorithm is very lucid and efficient to conclude security and analyze 4 bit S-Boxes. The algorithm can easily be expanded to 8 bit, 16 bit or 32 bit S-Boxes.

7. References:

1. Matsui, M. "The first experimental cryptanalysis of the data encryption standard". Advances in Cryptology - CRYPTO 1994.
2. Heys, H.M, "A Tutorial on Linear and Differential Cryptanalysis", Year:2001.Link: http://www.engr.mun.ca/~howard/PAPERS/ldc_tutorial.pdf.
3. Heys,H.M, Tavares,S.E, "Substitution-Permutation Networks Resistant to Differential and Linear Cryptanalysis", Jr. of Cryptology,vol.9,no.1,pp.1-19,1996.
4. Canniere, Christophe De, ""Analysis And Design Of Symmetric Encryption Algorithms", Mei: 2007, Link: http://image.sciencenet.cn/olddata/kexue.com.cn/upload/blog/file/2009/3/20093320521938772.pdf.
5. Data Encryption Standard, Federal Information Processing Standards Publication (FIPS PUB) 46, National Bureau of Standards, Washington, DC (1977).

6. Data Encryption Standard (DES), Federal Information Processing Standards Publication (FIPS PUB) 46-3, National Institute of Standards and Technology, Gaithersburg, MD (1999).
7. DES: S-oxesLink:http://www.gyyv.vd.ch/branches/mathematique/cryptographie/textes/s_boxes.htm.
8. Adams, Carlisle, Tavares, Stafford, "The structured design of cryptographically good S-boxe7s", J. Cryptology (1990) 3:27-41, Link: https://bitbucket.org/.../Sbox/The%20structured%20design%20of%20cry.

APPENDIX

In this section security analysis of 32 4-bit DES S-Boxes has been carried out. The analysis is demonstrated in Table.1. of Appendix.

4-Bit S-Box	Total Linear Equations	Security
E 4 D 1 2 F B 8 3 A 6 C 5 9 0 7	25	Good
0 F 7 4 E 2 D 1 A 6 C B 9 5 3 8	23	Medium
4 1 E 8 D 6 2 B F C 9 7 3 A 5 0	18	Medium
F C 8 2 4 9 1 7 5 B 3 E A 0 6 D	25	Good
F 1 8 E 6 B 3 4 9 7 2 D C 0 5 A	28	Good
3 D 4 7 F 2 8 E C 0 1 A 6 9 B 5	27	Good
0 E 7 B A 4 D 1 5 8 C 6 9 3 2 F	28	Good
D 8 A 1 3 F 4 2 B 6 7 C 0 5 E 9	28	Good
A 0 9 E 6 3 F 5 1 D C 7 B 4 2 8	17	Medium
D 7 0 9 3 4 6 A 2 8 5 E C B F 1	25	Good
D 6 4 9 8 F 3 0 B 1 2 C 5 A E 7	21	Medium
1 A D 0 6 9 8 7 4 F E 3 B 5 2 C	23	Medium
7 D E 3 0 6 9 A 1 2 8 5 B C 4 F	24	Good
D 8 B 5 6 F 0 3 4 7 2 C 1 A E 9	24	Good
A 6 9 0 C B 7 D F 1 3 E 5 2 8 4	22	Medium
3 F 0 6 A 1 D 8 9 4 5 B C 7 2 E	22	Medium
2 C 4 1 7 A B 6 8 5 3 F D 0 E 9	29	Good
E B 2 C 4 7 D 1 5 0 F A 3 9 8 6	25	Good
4 2 1 B A D 7 8 F 9 C 5 6 3 0 E	27	Good
B 8 C 7 1 E 2 D 6 F 0 9 A 4 5 3	20	Medium
C 1 A F 9 2 6 8 0 D 3 4 E 7 5 B	30	Good
A F 4 2 7 C 9 5 6 1 D E 0 B 3 8	30	Good
9 E F 5 2 8 C 3 7 0 4 A 1 D B 6	25	Good
4 3 2 C 9 5 F A B E 1 7 6 0 8 D	32	Good
4 B 2 E F 0 8 D 3 C 9 7 5 A 6 1	32	Good
D 0 B 7 4 9 1 A E 3 5 C 2 F 8 6	23	Medium
1 4 B D C 3 7 E A F 6 8 0 5 9 2	23	Medium
6 B D 8 1 4 A 7 9 5 0 F E 2 3 C	35	Good
D 2 8 4 6 F B 1 A 9 3 E 5 0 C 7	31	Good
1 F D 8 A 3 7 4 C 5 6 B 0 E 9 2	22	Medium
7 B 4 1 9 C E 2 0 6 A D F 3 5 8	14	Poor
2 1 E 7 4 A 8 D F C 9 0 3 5 6 B	23	Medium

Table.1. Security Analysis of 32 4-bit S-Boxes