Chapter 4: Diffusion Analysis in Block Cipher using SAC †

This chapter describes a scheme aimed at measuring the diffusion characteristic of a block cipher. The cryptographic strength of a cipher is directly proportional to the extent to which diffusion is achieved. The scheme described in this chapter is used to measure the above mentioned characteristic of a cipher and the test results are analyzed to come to a conclusion. Potentially, the scheme can be added as a part of an already existing test suite to act as a distinguisher based on the diffusion characteristic of the underlying cipher.

4.1 Introduction

Many test suites have been designed to test the extent of randomness approximated by a block cipher **[10].** Most of these tests measure the degree of randomness of change at block level by changing a bit in the original block **[18]**. However, while operating at block level, a situation may arise where the i^{th} bit of the block changes with a very high frequency whereas some other j^{th} bit hardly changes. This gives a false impression that all the bits are changing with a probability of 0.5.

In this chapter, a scheme is being proposed which is not significantly different in nature with some of the existing test sets but is rather different in terms of how it is implemented to measure the diffusion characteristic of the concerned cipher.

The scheme named "Bit-level Block Cipher Diffusion Analysis Test (BLDAT)" **[56]** is aimed at how vulnerable the underlying block cipher is with regards to a particular bit.

Like most tests in this field, the proposed scheme also treats the underlying block cipher as a black box and the results of analysis are based solely on the input to and output from the cipher under test.

4.2 Bit-Level Diffusion Analysis Test (BLDAT)

The scheme uses a randomly selected n bit block of plaintext (say P), which is then encrypted using the underlying cipher to produce the corresponding cipher block (say C) **[46]**. Then, a matrix of size $n \times n$ is produced, where each row of the matrix is P_i, a new plaintext block in itself derived from the original block by flipping the bit at the i^{th} position i.e. $P_i[i] = P \oplus e_i$, where e_i is a zero vector containing 1 at i^{th} position.

† This chapter is referenced from the published research paper: " A Bit-Level Block Cipher Diffusion Analysis Test, Springer International Publishing Switzerland 2015, Vol. 1, Advances in Intelligent Systems and Computing 327, DOI: 10.1007/978-3-319-11933-5_75".

$$P_i = \begin{pmatrix} p[0,0] & \cdots & p[0,n-1] \\ \vdots & \ddots & \vdots \\ p[n-1,0] & \cdots & p[n-1,n-1] \end{pmatrix}$$

Next, each row of the P_i matrix is fed as input to the underlying cipher to produce the corresponding ciphertext, which is stored as the i^{th} row of the $C_i$matrix of size $n \times n$.

$$C_i = \begin{pmatrix} c[0,0] & \cdots & c[0,n-1] \\ \vdots & \ddots & \vdots \\ c[n-1,0] & \cdots & c[n-1,n-1] \end{pmatrix}$$

i.e. $C_i[i] = E(P_i[i])$ where E() denotes encryption using the underlying block cipher.

At this point, the scheme kicks in to produce another matrix (say X) of size $n \times n$, where i^{th} row obtained by bitwise modulo 2 addition of C_i vector with C vector **[19]**, $X[i] = C_i[i] \oplus C$.

$$X = \begin{pmatrix} x[0,0] & \cdots & x[0,n-1] \\ \vdots & \ddots & \vdots \\ x[n-1,0] & \cdots & x[n-1,n-1] \end{pmatrix}$$

The scheme further produces the diffusion-factor by scanning each column of the X matrix. The algorithm of the proposed scheme is given in section 4.2.1.

4.2.1 Algorithm

Algorithm – BLDAT:

Step-1: Randomly select a binary string of n bits (P).

Step-2: Encrypt the plaintext with the concerned encryption algorithm to generate the corresponding ciphertext (C).

Step-3: Encrypt P_i's with the particular encryption algorithm to generate C_i's, where $P_i = P \oplus e_i$and e_i is a string of zeros with the i^{th} bit 1 and $E(P_i) = C_i$.

Step-4: $X_i = C_i \oplus C$ is stored in the i^{th} row of the matrix of size $n \times n$ where n, the number of bits is in the original plaintext.

Step-5: Find the number of $1s$ in each column (j).

As it is evident from the algorithm, there are n bits in the block and at every instance we had changed only one bit, as a result there are n blocks such that $H(P, P_i[i] = 1)$, where H denotes the Hamming Distance. The test finds, the number of times, a particular bit has changed when each of the n newly generated blocks are encrypted using the underlying cipher with respect to the ciphertext of the original block. Ideally each bit should change n/2 times, if a particular bit has changed with very high or very low frequency, it might motivate an attack.

The number of times a particular bit has changed is referred to as the vulnerability factor of the bit. An extremely low or extremely high vulnerability factor associated with a particular bit may act as a motivation for attackers to exploit this idea.

The time taken to construct the X matrix is $O(n^2)$ and the time taken to determine the bit-vulnerability factors is also $O(n^2)$. So, the time complexity of the algorithm is $O(n^2)$.

4.3 Experiment

A single block cipher has been used for applying the scheme. The ciphertext generated by the block cipher is subjected to the proposed test analysis.

4.3.1 Describing the Test Cipher

The analysis of the scheme is done using a simple Test Cipher. The Test Cipher is a simple substitution – permutation network which takes 8-bit block as input and produces 8-bit cipher block. At first, as found in **[4]**, the 8 bit block is bit wise XOR-ed with an 8-bit key and then passed through two 4-bit S-boxes (for simplicity the two S-boxes are considered to be identical). The outputs of the S-boxes are permuted to generate the ciphertext. The block diagram of the test cipher is depicted by Figure 4.1.

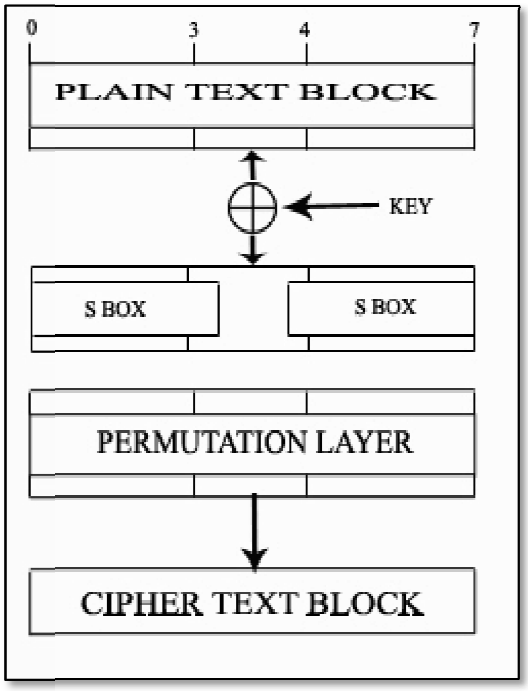

Figure 4.1: Block diagram of test cipher

4.3.2 Objective

The objective is to determine the secret key as is the case with other cryptanalysis techniques.

4.3.3 Assumptions

It is assumed that the plaintext block is known, the corresponding ciphertext block is also known and the results of the bit-level diffusion test are available. Using exhaustive key search method, the correct key can be determined with 28 trials (in the worst case). The sub-goal would be to reduce the number of trials using the results of BLDAT. If it is observed from the results of the BLDAT that a particular bit (say j^{th} bit) of the cipher block seldom changes i.e. if the i^{th} bit of the plaintext is 1/0, and remains 1/0 in the ciphertext at j^{th} position with a very high degree of probability, then it will be easy to identify the particular key bit. Linear cryptanalysis and differential cryptanalysis are well known for mapping an input plaintext bit to an output ciphertext bit. If the observed plaintext and ciphertext are dissimilar, then it is clearly due to the key bit which got XOR-ed with the i^{th} plaintext bit and will be a 1with a very degree of probability. And if both the observed bits are the same, then it implies that the key bit has not affected the plaintext bit which in turn implies that the bit is a 0 with a very high degree of probability.

4.3.4 Experimental Results of BLDAT

Two well-known ciphers, namely Data Encryption Standard and Rijndael (Advanced Encryption Standard) are put to test, and the results obtained are analyzed in **[1]**. Say δ is deviation where $\delta = 0$ is the ideal case. Table 4.1 and Table 4.2 lists the results of BLDAT on DES and Rijndael Cipher (AES) respectively.

Key	No. of Deviations			
	$\boldsymbol{\delta = 0}$	$\boldsymbol{\delta = 4}$	$\boldsymbol{\delta = 8}$	$\boldsymbol{\delta > 8}$
Sparse Key	5	40	16	3
Moderately Dense Key	5	47	9	3
Dense Key	5	43	13	3
Random Key	7	34	21	2

Table 4.1. Experimental result on DES

Key	No. of Deviations			
	$\boldsymbol{\delta = 0}$	$\boldsymbol{\delta = 8}$	$\boldsymbol{\delta = 16}$	$\boldsymbol{\delta > 16}$
Sparse Key	8	89	13	1
Moderately Dense Key	11	99	18	0
Dense Key	5	102	20	1
Random Key	6	97	25	0

Table 4.2. Experimental result on AES

4.4 Analysis of BLDAT

The results obtained from both DES and AES are analyzed using the established statistical tools to finally draw the conclusion. The Chi-square (χ^2) test has been used to determine the goodness of fit between theoretical and experimental data. The observed values and expected values are to be tested here.

4.4.1 Chi- Square (χ^2) test on experimental result of DES

In the experiment of BLDAT for DES, with 64 bit plaintext block and 56 bit key, the observed bit changes in ciphertext block for every bit change in plaintext given in Table 4.3.

v[0]=29	v[8]=27	v[16]=30	v[24]=38	v[32]=32	v[40]=37	v[48]=37	v[56]=24
v[1]=38	v[9]=37	v[17]=36	v[25]=33	v[33]=32	v[41]=28	v[49]=30	v[57]=29
v[2]=31	v[10]=31	v[18]=30	v[26]=38	v[34]=30	v[42]=25	v[50]=33	v[58]=35
v[3]=35	v[11]=32	v[19]=28	v[27]=34	v[35]=33	v[43]=32	v[51]=24	v[59]=26
v[4]=26	v[12]=33	v[20]=24	v[28]=28	v[36]=31	v[44]=31	v[52]=24	v[60]=23
v[5]=28	v[13]=29	v[21]=34	v[29]=31	v[37]=34	v[45]=33	v[53]=40	v[61]=27
v[6]=28	v[14]=38	v[22]=39	v[30]=30	v[38]=36	v[46]=31	v[54]=35	v[62]=30
v[7]=26	v[15]=32	v[23]=32	v[31]=31	v[39]=32	v[47]=42	v[55]=31	v[63]=33

Table 4.3. Observed bit changes in ciphertext using DES

The change of ciphertext block is ideally *n/2* where n is the block size and is estimated 32. Table 4.4 is constructed to calculate the Chi-square distribution and goodness of fit for Chi-square.

(T)	(O)	(E)	O-E	$(O-E)^2$	$Y=(O-E)^2/E$
3	29	32	-3	9	0.28125
4	38	32	6	36	1.125
8	31	32	-1	1	0.03125
3	35	32	3	9	0.28125
3	26	32	-6	36	1.125
5	28	32	-4	16	0.5
2	27	32	-5	25	0.78125
3	37	32	5	25	0.78125
6	33	32	1	1	0.03125
7	32	32	0	0	0
6	30	32	-2	4	0.125
2	36	32	4	16	0.5
4	24	32	-8	64	2
3	34	32	2	4	0.125
1	39	32	7	49	1.53125
1	25	32	-7	49	1.53125
1	42	32	10	100	3.125
1	40	32	8	64	2
1	23	32	-9	81	2.53125

Table 4.4. DES observed values with corresponds to estimated value with their occurrence

Where T is occurrences of observed value, O is observed value, E is estimated value. Using these values the chi-square is calculated with the consideration of occurrence of values and the calculated chi-square value of the experimental data is:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$\chi^2 = 37.25$$

and Figure. 4.2 is the graphical representation of observed value, estimated value and calculated value:

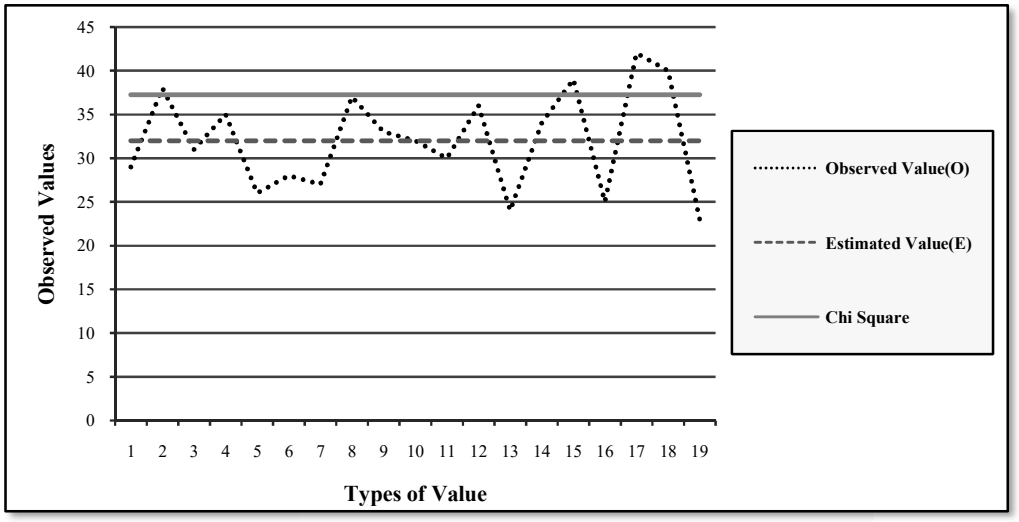

Figure 4.2: Graph for Chi-square of DES

In the experimental result (Table 4.4) it is visible that there exist 19 different sets of value (v) i.e. number of time changes of every bit in ciphertext while changing every bit of plaintext for at most once. Therefore the Degree of Freedom (*df*) is:

$$df = v - 1 = 19 - 1 = 18$$

From the Table 4.4 it is now easy to calculate chi-square value for $df = 18\ and\ \alpha = .005$ is:

$$GF_{18,.005} = \chi^2_{18,.005} = 34.71875$$

Moreover, chi-square value for $df = 18\ and\ \alpha = 5\%$ is:

$$GF_{18,5\%} = \chi^2_{18,5\%} = 31.5$$

From the chi-square distribution table **[20]**, is found that $\chi^2_{18,.005}$ is 37.156 and $GF_{18,.005} < 37.156$.

So, according to **[21]** it may be concluded that either (i) this model is valid but that a statistically improbable excursion of χ^2 has occurred, (ii) too conservatively, over-estimated the values of α or (iii) data is 'too good to be true'.

4.4.2 Chi- Square (χ^2) test on experimental result of AES

In the experiment of BLDAT for AES, for 128 bit plaintext block with 56 bit key and observed bit changes in ciphertext block for every bit change in plaintext. The Chi-square (χ^2) test has been used. The change of ciphertext block is ideally $n/2$ where n is the block size and is estimated 64. The calculated chi-square value of the experimental data of AES is:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$\chi^2 = 73.609375$$

Fig. 4.3 is the graphical representation of observed value, estimated value and calculated value.

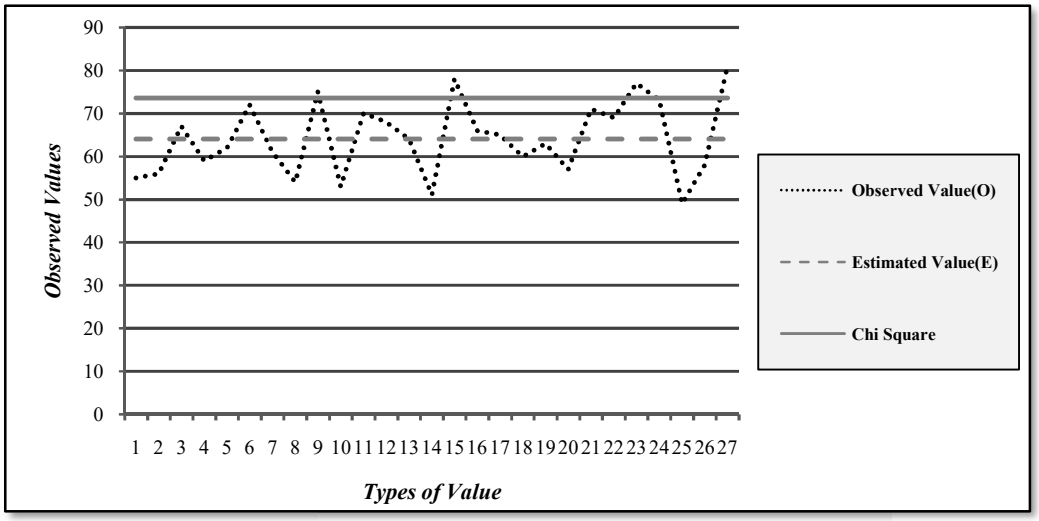

Figure 4.3: Graph for Chi-square of AES

In the experimental it is visible that there exist 27 different sets of value (v) i.e. number of time changes of every bit in ciphertext while changing every bit of plaintext for at most once. Therefore the Degree of Freedom (*df*) is:

$$df = v - 1 = 27 - 1 = 26$$

Now it is easy to calculate chi-square value for $df = 26 \; and \; \alpha = .005$ is:

$$GF_{26,.005} = \chi^2_{26,.005} = 69.09375$$

Moreover, chi-square value for $df = 26 \; and \; \alpha = 5\%$ is:

$$GF_{26,5\%} = \chi^2_{26,5\%} = 63.8046875$$

From the chi-square distribution table **[20]**, is found that $\chi^2_{26,.005}$ is 48.290 and $GF_{26,.005} > 48.290$.

Therefore, it may be concluded that either (i) this model is valid one but that a statistically improbable excursion of χ^2 has occurred or (ii) that this model is poorly chosen that an unacceptable large value of χ^2 has resulted **[21]**. The theory of chi-square test relies on the assumption that chi-square is the sum of the squares of random normal derivatives, that is, that each x_i is normally distributed over its mean value μ_i.

4.5 Conclusion on BLDAT

Even if the Hamming Distance of the plaintext block and the ciphertext block is ideal i.e., $n/2$, where n is the block size, the key space can be reduced using a scheme such as the proposed BLDAT. The scheme may be clubbed with other tests for comprehensive analysis of block ciphers.