Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2023.020513
Article

A Novel Approach for Security Enhancement of Data Encryption Standard

Dawood Shah[1,*], Tariq Shah[1], Sajjad Shaukat Jamal[2], Mohammad Mazyad Hazzazi[2], Amer Aljaedi[3] and Adel R. Alharbi[3]

[1]Department of Mathematics, Quaid-i-Azam University, Islamabad, Pakistan
[2]Department of Mathematics, College of Science, King Khalid University, Abha, Saudi Arabia
[3]College of Computing and Information Technology, University of Tabuk, Tabuk, 71491, Saudi Arabia

*Corresponding Author: Dawood Shah. Email: dawoodshah@math.qau.edu.pk
Received: 27 May 2021; Accepted: 14 January 2022

Abstract: Data Encryption Standard (DES) is a symmetric key cryptosystem that is applied in different cryptosystems of recent times. However, researchers found defects in the main assembling of the DES and declared it insecure against linear and differential cryptanalysis. In this paper, we have studied the faults and made improvements in their internal structure and get the new algorithm for Improved DES. The improvement is being made in the substitution step, which is the only nonlinear component of the algorithm. This alteration provided us with great outcomes and increase the strength of DES. Accordingly, a novel 6×6 good quality S-box construction scheme has been hired in the substitution phase of the DES. The construction involves the Galois field method and generates robust S-boxes that are used to secure the scheme against linear and differential attacks. Then again, the key space of the improved DES has been enhanced against the brute force attack. The outcomes of different performance analyses depict the strength of our proposed substitution boxes which also guarantees the strength of the overall DES.

Keywords: DES; S-box; linear cryptanalysis; differential cryptanalysis

1 Introduction

With rapid development in the field of information technology, it is observed that communication over electronics channels and broadcasting of digital data over the internet is increased. Accordingly, the security of sensitive information against prohibited copying and dissemination has become tremendously imperative. Cryptography is the study of secure communication which is contemplated as a recognized branch of science for the last 60 years. However, it is quite a new area of study comparable to other areas of sciences as each moment carries continual developments. Cryptography is divided into two sub-branches; asymmetric key cryptography, and symmetric key cryptography. This classification is based on the secret key that is used during encryption and decryption. In symmetric-key cryptography, the communicating parties share a secret key confidentially. The algorithms such as Lucifer, DES, Advanced encryption standard (AES), and the International data encryption algorithm (IDEA) are the prominent examples of symmetric key cryptography.

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data encryption standard (DES) is a symmetric key algorithm designed by IBM. It was adopted and published by the US National Institute of Standard Technology (NIST) in 1971, as a federal information processing scheme. The aim was to provide a secure cryptosystem for the security of sensitive data and information during transmission. This algorithm became a distinguished and broadly used algorithm [1]. In the same way, a considerable number of cryptanalytic papers on DES were published since its acceptance in 1971. In 1977, Diffie and Hellman proposed a parallel machine for the comprehensive search of the complete keyspace [2]. The author claimed was, that very-large-scale integration (VLSI) chips are constructed and each chip probably searches one key per microsecond. The construction of the search machine contains millions of such chips and all working in parallel mode. Each chip is credible to search 1012 keys per second. The order of the set of all keys of DES is 7×10^{16} and can be approximately searched in 10^5 seconds that is almost 24 h. The estimated cost of this machine was $ 20 million and hence the cost per solution was $ 5000. In 1980 Hellman presented a time-memory tradeoff technique for the chosen plaintext attack [3]. The time memory tradeoff method takes vu words of memory and performed u^2 operations. Since vu^2 equivalent to the total number of all possible keys for DES. Therefore, this technique is the same as the Differential Cryptanalysis of DES-like Cryptosystems which takes about 2^{38} times and 2^{38} memory with 256 pre-processing time for a special case $m = t$. The author suggested a special machine that produced about a hundred solutions with an average time of 24 h. The approximate cost of that machine was $ 4 million, so the cost per solution was about $ 1–$ 100. The processing time for the same machine was estimated and was required two years or three years. In 1985, Evertse and Chaum showed that the meet-in-the-middle attack can decrease the key search for DES [4]. The reduction factors are 2^{19}, 2^9 and 2^2 for the reduced number of rounds 4, 5 and 6 respectively. They also claimed that a somewhat altered form of DES having seven rounds can be cracked through the reduction factor of 2. Besides this, they showed that a meet-in-the-middle attack of the same kind is not appropriate for eight or more round reduced DES. In 1987, Davies described a new kind of cryptanalytic attack on DES called it known-plaintext attack [5]. They assumed that sufficient data might produce sixteen linear relationships amid the key bits. Accordingly, it decreases the size of the key search up to 2^{40}. The correlation among the outputs of the adjacent S-boxes was the main target of the plaintext attack. Since the correlation can disclose the linear relationship between the four bits of the key that is utilized to adjust these S-boxes as an input bit. Moreover, the consequence of the splits 32-bit of DES receives these outputs independently. Thus, each pair of the adjacent S-boxes can be exploited twofold, yielding 16 bits of key information. In 1991, Eli Biham and Adil Shamir designed the differential attacks which applied to various DES-like substitution permutation cryptosystems [5]. This was a powerful attack, which used just the pairs of ciphertexts and broke the DES in a few minutes. According to [6], any modification in the algorithm, for instance, key scheduling of the algorithm, substituting the permutation step by any other permutation, or the change the order of the eight S-boxes cannot make the algorithm less successful against the differential attack. A complete review of these attacks shows that the main targets of this cryptanalysis are the substitution phase which is the only nonlinear part of the algorithm. Since the S-boxes used in the algorithm were not cryptographically strong and thus the DES proved to be insecure against differential attacks.

S-box is one of the most important components in block ciphers, which is used to confuse the relationship between the cipher data and the input key during the process of encryption. Since the confusion-creating capability of the cipher relies on the quality of the S-box. Therefore, the construction of good quality S-box has attracted the cryptographic research community. The S-box construction schemes based on all isomorphic Galois field $GF(2^8)$ is given in [6]. In [7], the author presented a novel S-box construction scheme using a chaotic skew tent map and its image encryption

application. The nonlinear dynamical system exhibits desirable properties that are useful for confusion, so these are widely used for S-box construction [8–10].

Keeping the above facts in view, this manuscript proposed a novel 6×6 cryptographically strong S-boxes. The proposed S-boxes are then deployed in the Feistel function (FF) of the DES to achieve the substitution transformation which is the necessary step for the confusion criterion. The cryptographic characteristics of the new S-boxes are then analyzed over different analyses such as Differential approximation probability (DP), linear approximation probability (LP), Nonlinearity, strict avalanche criterion (SAC) and bit independent criterion (BIC). The results show that the proposed S-boxes are bijective, highly nonlinear and low costly than AES 8-bit S-box to implement and are almost identical in the term of linear, differential and algebraic properties. Hence, the essential criterion for the substitution step of the DES algorithm is successfully achieved. The major contribution of this paper is to strengthen the DES algorithm against linear and differential attacks.

The rest of this manuscript is organized as follows. In Section 2, we introduced some basic definitions. Section 3 is dedicated to the general principle of DES. The construction of the proposed S-box construction scheme is presented in Section 4. Section 5 is devoted to the performance analyses of the suggested S-boxes. The modified DES is examined using a specific ciphertext and compared the results with the DES in Section 6. Section 7 concluded the discussion.

2 Preliminaries

We denote the direct product of n copies of the field $\mathbb{Z}_p$ by $\mathbb{Z}_p^n$ where p is a positive prime integer. The *2*-ary function having a range in $\mathbb{Z}_2$ is denoted by f throughout in this study namely Boolean functions, which is defined as $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$. However, the function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$ is called a vectorial Boolean function.

2.1 Definition

Let $f{:}\mathbb{Z}_2^n \to \mathbb{Z}_2$ be a Boolean function. Then the nonlinearity of f can be defined as the smallest hamming distance among the set of all affine Boolean functions and the function f. The nonlinearity of the function f is denoted by N_f. Mathematically it can be written as;

$$N_f = min\{d(f, a) : \ a \in A \} \tag{1}$$

where $d(f, a)$ denote the Hamming distance between f and a. The symbol A signifies the set of all affine Boolean functions. Accordingly, the maximum possible N_f value of a function f is equal to $2^{n-1} - 2^{\frac{n}{2}-1}$.

2.2 Definition

Followed by [11], a function $F{:}\mathbb{Z}_p^m \to \mathbb{Z}_p^n$ is said to exhibit the avalanche effect if and only if the following equation holds.

$$\sum_{y \in \mathbb{Z}_p^m} wt(F(y) \oplus F(X \oplus C_i^m)) = n \cdot 2^{m-1}. \tag{2}$$

For all i $(1 \leq i \leq m)$, Eq. (2), implies that the average of one-half of the output bits must be changed whenever one bit is complemented by the input data.

2.3 Definition

Followed by [11], let $F: \mathbb{Z}_p^m \to \mathbb{Z}_p^n$ be a function and defined a set

$$\partial(x, y) = |\{x \mid F(z + x) - F(z) = y\}| \tag{3}$$

The positive integer $\triangle_F$ is called differential $\triangle_F-$ uniform. The mathematical representation is defined as:

$$\triangle_F = \max_{\substack{x \in \mathbb{Z}_p{}^m, \quad x \neq 0 \\ y \in \mathbb{Z}_p{}^n}} \partial(x, y) \tag{4}$$

3 General Outline of DES

DES is a symmetric key encryption scheme, which encrypts a 64-bits block of data. Thus, the input size of the algorithm is 64-bits and the output size is also 64-bits. The length of the key is 56-bits and it is mostly expressed as the block of 64-bits. The 56-bits are used as input key and the remaining eight least significant bits are utilizes for the parity check purpose. DES is consisting of two modules that are product cipher and Feistel cipher. The product cipher is used to combines two or more transformations because the combinations of ciphers are more secure than separated ciphers. A Feistel cipher is the iterated cipher that consists of the sequential repetition of the round functions. The formal definition of the FF is given as follows.

The Feistel Function (FF) is an iterated cipher that maps plaintext of size $n = 2m$. We denote the left t-bits block and right t-bits block of the plaintext by $\mathcal{L}_0$ and $\Re_0$ respectively. Assume that the FF is consist of r rounds and the output of the r_{th} round is the ciphertext, thus we denote the ciphertext by $(\mathcal{L}_r, \Re_r)$. The Feistel for i_{th} round (for $1 \leq i \leq r$) is defined as follows;

$$(\mathcal{L}_{i-1}, \Re_{i-1}) \mapsto (\mathcal{L}_i, \Re_i) \tag{5}$$

$$(\mathcal{L}_i, \Re_i) = \begin{cases} \mathcal{L}_i = \Re_{i-1} \\ \Re_i = \mathcal{L}_{i-1} \oplus f(\Re_{i-1}, \mathcal{K}_i) \end{cases} \tag{6}$$

where $\mathcal{K}_i$ is the subkey derived through the key schedule algorithm. In DES the number of rounds $r = 16$ and the subkeys size $\mathcal{K}_i$ is 48-bits.

The FF is bijective and thus reversible. So, the same key is used for the encryption and decryption procedure. The Xor is used in the function to combine the output of the round function with the left half using the following equation.

$$\mathcal{L}_{i-1} \oplus f(\Re_{i-1}, \mathcal{K}_i) \oplus f(\Re_{i-1}, \mathcal{K}_i) = \mathcal{L}_{i-1} \tag{7}$$

Eq. (7), demonstrates that the DES algorithm is independent of the design of the FF. The invertibility of FF does not produce an impact on the invertibility of the DES algorithm. Accordingly, if the function FF is invertible or not, though the DES scheme is always invertible.

4 Construction of Galois Fields $GF(2^6)$ and S-Boxes

The field of order p is a prime field that is denoted by $\mathbb{Z}_p$. A polynomial $p(y) \in \mathbb{Z}_p[y]$ that cannot factor in the product of polynomials in $\mathbb{Z}_p[y]$ is called irreducible polynomials. Let $p(y)$ be an irreducible polynomial in principle ideal domain $\mathbb{Z}_p[y]$,. Therefore, the ideal generated by $p(y)$ is a

maximal ideal in $\mathbb{Z}_p[y]$. The ideal generated by $p(y)$ is denoted by $\langle p(y)\rangle$ and it is defined as;

$$\langle p(y)\rangle = \{a(y) : a(y) = p(y).h(y),\ for\ some\ h(y) \in \mathbb{Z}_p[y]\ \}. \tag{8}$$

Thus, the quotient ring $\frac{\mathbb{Z}_p[y]}{\langle p(y)\rangle}$ is a finite field of order p^n, which is known as Galois field $GF(p^n)$, where m is the degree of the polynomial $p(y)$. The field $\frac{\mathbb{Z}_p[y]}{\langle p(y)\rangle}$ consists of all polynomials in the form of the element of the principal ideal domain $\mathbb{Z}_p[y]$ having a degree strictly less than m. The subtraction and addition operations perform over the field $\mathbb{Z}_p$, that are the same operations as performed in $\mathbb{Z}_p[y]$. However, the product of the polynomials performs modulo $p(y)$. A polynomial $f(y) \in \frac{\mathbb{Z}_p[y]}{\langle p(y)\rangle}$ is said to be the multiplicative inverse of the non-zero polynomial $g \in \frac{\mathbb{Z}_p[y]}{\langle p(y)\rangle}$, if and only if $f(y)g(y) \equiv 1\ mod\ p(y)$.

4.1 Construction of Galois Fields

This main interest of this study is Galois field s $GF(2^6)$ of order 2^6. To construct the Galois field $GF(2^6)$, initially choose a degree 6 primitive irreducible polynomial $p(y)$ in $\mathbb{Z}_2[y]$ and find the root β of the polynomial $p(y)$ i.e., $p(\beta) = 0$. Subsequently, generate the multiplicative cyclic group $GF(2^6) - \{0\}$ from the root β by computing all β^i for $1 \le i \le 2^6 - 1$. Hence each nonzero element of the field $GF(2^6)$ can be expressed as the power of the primitive element β. We consider the set $\{p_i(y) \in \mathbb{Z}_2[y] : p_i(y)$ is irreducible and $1 \le i \le 6\}$ of all primitive irreducible polynomials of degree 6, to construct corresponding the Galois Fields $\frac{\mathbb{Z}_p[y]}{\langle p_i(y)\rangle}$, $1 \le i \le 6$. Next, these Galois fields are then utilized to construct 6×6 S-boxes. The degree 6 primitive irreducible polynomials and their corresponding Galois fields are listed in Table 1.

Table 1: List of degree 6 primitive irreducible polynomials over $\mathbb{Z}_2$

Primitive polynomials	$GF(2^6)$	Primitive polynomials	$GF(2^6)$
$p_1(y) = y^6 + y + 1;\ \beta_1$	$\frac{\mathbb{Z}_2[y]}{\langle p_1(y)\rangle}$	$p_4(y) = y^6 + y^5 + 1;\ \beta_4$	$\frac{\mathbb{Z}_2[y]}{\langle p_4(y)\rangle}$
$p_2(x) = y^6 + y^4 + y^3 + y + 1;\ \beta_2$	$\frac{\mathbb{Z}_2[y]}{\langle p_2(y)\rangle}$	$p_5(x) = y^6 + y^5 + y^3 + y^2 + 1;\ \beta_5$	$\frac{\mathbb{Z}_2[y]}{\langle p_5(y)\rangle}$
$p_3(x) = y^6 + y^5 + y^2 + y + 1;\ \beta_3$	$\frac{\mathbb{Z}_2[y]}{\langle p_3(y)\rangle}$	$p_3(x) = y^6 + y^5 + y^4 + y + 1;\ \beta_3$	$\frac{\mathbb{Z}_2[y]}{\langle p_6(y)\rangle}$

4.2 Construction of 6×6 *S-Boxes*

The construction of the S-box required a nonlinear bijective map. In the proposed work, we have used the multiplicative inverse function module degree 6 primitive irreducible polynomial $p_i(y)$ as a power permutation for the construction of S-boxes. The mapping is defined as follows:

$$g_i : \frac{\mathbb{Z}_2[y]}{\langle p_i(y)\rangle} \to \frac{\mathbb{Z}_2[y]}{\langle p_i(y)\rangle}$$

$$g_i(w) = \begin{cases} w^{-1} & if\ w \neq 0 \\ 0 & if\ w = 0 \end{cases} \tag{9}$$

The images $g_i(w)$ for all $0 \le w \le 63$ are then converted into an 8×8 lookup table, which is the required S-box. Thus, for each degree 6 primitive irreducible $p_i(w)$ for $1 \le i \le 6$ one can obtain

different S-box denoted by S_i. Tables 2a–2f depicted the generated S-boxes corresponding to different primitive irreducible polynomials and Galois field $\frac{\mathbb{Z}_2[y]}{\langle p_i(y)\rangle}$. Section 5, analyzed the proposed S-boxes with well-known analyses such as nonlinearity, BIC, SAC, LP and DP to examine the quality of the S-boxes.

Table 2: Proposed S-boxes

(a) Proposed S-box 1 $\mathbf{S}_1$

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

(b) Proposed S-box 2 $\mathbf{S}_2$

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

(c) Proposed S-box 3 $\mathbf{S}_3$

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

(d) Proposed S-box 4 $\mathbf{S}_4$

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

(e) Proposed S-box 5 $\mathbf{S}_5$

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

(f) Proposed S-box 6 $\mathbf{S}_6$

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

4.3 Theorem

In [12], Let l be an affine transformation and g be the power permutation with good cryptographic properties in the Galois field $GF(2^m)$, then the affine power affine (APA) composition is defined as follows:

$$S(x) = l \circ g \circ l \tag{10}$$

The Eq. (10) preserves the cryptographic properties of g and takes on stronger algebraic complexity.

4.4 Remark

A composition function of an affine function with a function g on the right-hand side or the left-hand side preserved the properties of linearity and differential uniformity of a function g.

5 Performance Analyses

An efficient cryptosystem should be secure against all kinds of attacks. Since the security of the block ciphers depends on the choice of the S-box, therefore this section we thoroughly analyzed the performance of the proposed 6×6 S-boxes to figure out the best S-box. The good quality S-box of these S-boxes is then deployed in the proposed modified DES. Besides, we will also compare the obtained results with the super AES 8-bits S-box.

5.1 Nonlinearity

In Section 2, the definition of nonlinearity for the Boolean function has been already discussed. The general formula to calculate the upper bound of the nonlinearity of the function $f : \mathbb{Z}_2^m \to \mathbb{Z}_2$ that is $2^{m-1} - 2^{\frac{m}{2}-1}$ [13]. Therefore, the maximum possible nonlinearity for $m = 6$ is $N_{max} = 28$. The nonlinearity of all S-boxes is calculated, the resultant values are listed in Tables 3a–3f. It can be seen that overall, the average nonlinearity analyses of the S-boxes are quite good and capable to resist linear attacks. Moreover, from the tables, one can observe that the average nonlinearity of the S-box S_1 is equal to S_2. Similarly, the average nonlinearity value of S_3 is equal to S_4 and the S_5 average nonlinearity value is the same as S_6. Therefore, the pair consists of S_5 and S_6 is the best with respect to nonlinearity analysis.

Table 3: Nonlinearity of the proposed S-boxes

(a). The nonlinearity of the S-box $\boldsymbol{S}_1$

Function	f_0	f_1	f_2	f_3	f_4	f_5
Nonlinearity	24	22	16	24	20	20

(b) The nonlinearity of the S-box $\boldsymbol{S}_2$

Function	f_0	f_1	f_2	f_3	f_4	f_5
Nonlinearity	22	22	22	22	22	20

(c) The nonlinearity of the S-box $\boldsymbol{S}_3$

Function	f_0	f_1	f_2	f_3	f_4	f_5
Nonlinearity	20	24	22	20	22	22

(d) The nonlinearity of the S-box $\boldsymbol{S}_4$

Function	f_0	f_1	f_2	f_3	f_4	f_5
Nonlinearity	22	24	22	22	22	20

(e) The nonlinearity of the S-box $\boldsymbol{S}_5$

Function	f_0	f_1	f_2	f_3	f_4	f_5
Nonlinearity	24	22	22	22	24	20

(f) The nonlinearity of the S-box $\boldsymbol{S}_6$

Function	f_0	f_1	f_2	f_3	f_4	f_5
Nonlinearity	24	20	22	22	24	22

5.2 Differential Cryptanalysis

Differential approximation probability (DP) analysis is used to measure the differential uniformity of the S-box. The minimum possible value of differential uniformity for $m \times n$ S-box is

$\delta(S) = 2^{m-n+1}$ [14]. Thus, for the 6-bit S-box, $m = n = 6$, $\delta_{min} = 2$. The S-box having minimum differential uniformity is known as almost perfect nonlinear [15]. We have calculated the differential distribution matrix Λ(S) of all the generated S-boxes, shown in Tables 4a–4f. As can be seen in the tables that the differential distribution table of all the 6-bit S-boxes are consist of 4, 6 and 8 except the element λ_{63} and the S-box S_5 also contain 10. Therefore, the differential approximation probability is 0.1250 for the S-boxes $S_1, \ldots, S_4$ and S_6, however the differential probability of the S-box S_5 is 0.1563. Overall, the differential approximation values of all S-box are approximately equal to the DP value of the AES S-box, nowadays considered as a super S-box. Accordingly, the modified DES S-boxes have enough strength against the differential cryptanalysis attack.

Table 4: DP analysis of the Proposed S-boxes

(a) DP table of $\boldsymbol{S}_1$

6	6	8	4	4	6	6	4	6	4	6	6	6	4	4	4	6	4	6	6	6	4	4	4
4	6	4	6	6	4	6	4	6	6	4	8	8	6	4	6	6	6	8	6	6	4	4	4
4	6	4	6	6	4	6	6	6	6	8	6	6	4	6	6	4	4	6	4	4	4	6	6
6	6	4	6	4	6	6	6	6	6	4	4	6	4	6	8	8	6	4	4	6	4	8	6
6	6	6	8	4	4	6	4	6	4	6	6	4	4	4	6	4	6	4	4	8	6	4	8
4	4	4	4	4	8	4	4	8	4	6	4	6	6	6	4	6	8	4	4	6	6	6	6
8	8	6	6	6	6	8	6	4	6	4	6	6	4	4	8	6	6	4	4	4	4	6	4
6	4	4	8	4	4	6	0	4	6	4	6	8	6	6	0	4	6	6	6	6	6	6	0

(b) DP table of $\boldsymbol{S}_2$

4	6	6	8	6	4	6	6	6	6	6	4	4	6	4	6	6	6	6	6	6	6	6	6
4	8	6	6	4	6	6	4	6	4	4	4	8	4	6	6	4	6	8	6	6	6	4	6
6	8	6	4	6	4	6	4	4	6	4	4	4	4	6	6	8	4	4	6	4	4	4	8
8	4	6	4	4	6	6	8	4	4	4	4	6	6	6	6	4	4	6	6	6	4	6	4
6	6	8	6	4	8	4	4	4	4	8	4	6	4	8	6	6	4	6	4	6	6	6	4
8	4	8	4	6	6	8	6	4	6	6	6	6	4	6	4	6	6	6	6	6	4	6	6
6	8	6	6	6	4	6	8	4	4	6	4	4	4	8	4	4	6	4	4	4	4	6	4
6	8	4	8	6	6	6	0	4	1	4	4	4	6	6	0	8	6	6	8	6	6	4	0

(c) DP table of $\boldsymbol{S}_3$

6	4	6	6	6	4	4	4	6	6	4	6	6	6	4	4	4	6	4	6	6	6	4	4	4
4	6	4	6	6	4	6	4	6	6	4	8	8	6	4	6	6	6	8	6	6	4	4	4
4	6	4	6	6	4	6	6	6	6	8	6	6	4	6	6	4	4	6	4	4	4	6	6
6	6	4	6	4	6	6	6	6	6	4	4	6	4	6	8	8	6	4	4	6	4	8	6
6	6	6	8	4	4	6	4	6	4	6	6	4	4	4	6	4	6	4	4	8	6	4	8
4	4	4	4	4	8	4	4	8	4	6	4	6	6	6	4	6	8	4	4	6	6	6	6
8	8	6	6	6	6	8	6	4	6	4	6	6	4	4	8	6	6	4	4	4	4	6	4
6	4	4	8	4	4	6	0	4	6	4	6	8	6	6	0	4	6	6	6	6	6	6	0

(d) DP table of $\boldsymbol{S}_4$

4	6	6	8	6	4	6	6	6	6	6	4	4	6	4	6	6	6	6	6	6	6	6	6
4	8	6	6	4	6	6	4	6	4	4	4	8	4	6	6	4	6	8	6	6	6	4	6
6	8	6	4	6	4	6	4	4	6	4	4	4	4	6	6	8	4	4	6	4	4	4	8
8	4	6	4	4	6	6	8	4	4	4	4	6	6	6	6	4	4	6	6	6	4	6	4
6	6	8	6	4	8	4	4	4	4	8	4	6	4	8	6	6	4	6	4	6	6	6	4
8	4	8	4	6	6	8	6	4	6	6	6	6	4	6	4	6	6	6	6	6	4	6	6
6	8	6	6	6	4	6	8	4	4	6	4	4	4	8	4	4	6	4	4	4	4	6	4
6	8	4	8	6	6	6	0	4	1	4	4	4	6	6	0	8	6	6	8	6	6	4	0

(e) DP table of $\boldsymbol{S}_5$

6	6	6	6	6	6	6	6	6	6	6	4	4	6	4	6	6	6	6	6	6	6	6	6
4	6	6	8	6	4	6	6	6	6	6	4	4	6	4	6	6	6	6	6	6	6	6	6
6	6	6	4	6	4	6	4	4	6	4	4	4	4	6	6	8	4	4	6	4	4	4	8
6	4	6	4	4	6	6	8	4	4	4	4	6	6	6	6	4	4	6	6	6	4	6	4
6	6	8	6	4	8	4	4	4	4	8	4	6	4	8	6	6	4	6	4	6	6	6	4
8	4	8	4	6	6	8	6	4	6	6	6	6	4	6	4	6	6	6	6	6	4	6	6
6	8	6	6	6	4	6	8	4	4	6	4	4	4	8	4	4	6	4	4	4	4	6	4
6	8	4	8	6	6	6	0	4	1	4	4	4	6	6	0	8	6	6	8	6	6	4	0

(f) DP table of $\boldsymbol{S}_6$

6	6	6	6	6	6	6	6	6	6	6	4	4	6	4	6	6	6	6	6	6	6	6	6
4	6	6	8	6	4	6	6	6	6	6	4	4	6	4	6	6	6	6	6	6	6	6	6
6	6	6	4	6	4	6	4	4	6	4	4	4	4	6	6	8	4	4	6	4	4	4	8
6	4	6	4	4	6	6	8	4	4	4	4	6	6	6	6	4	4	6	6	6	4	6	4
6	6	8	6	4	8	4	4	4	4	8	4	6	4	8	6	6	4	6	4	6	6	6	4
8	4	8	4	6	6	8	6	4	6	6	6	6	4	6	4	6	6	6	6	6	4	6	6
6	8	6	6	6	4	6	8	4	4	6	4	4	4	8	4	4	6	4	4	4	4	6	4
6	8	4	8	6	6	6	0	4	1	4	4	4	6	6	0	8	6	6	8	6	6	4	0

5.3 Strict Avalanche Criterion

In general, an S-box is considered a lookup table of Boolean functions from $\mathbb{Z}_2^m$ to $\mathbb{Z}_2^n$ for $m \geq n$ [16]. Feistel has suggested an important criterion for the designation of cryptographic function. A Boolean function $f : \mathbb{Z}_2^m \to \mathbb{Z}_2^n$ is said to be exhibit the avalanche effect if

$$\sum_{u \in \mathbb{Z}_2^m} wt(f(u) \oplus f(u \oplus c_i^n)) = n2^{m-1}. \tag{11}$$

For all $1 \leq i \leq m$, where c_i^n is a vector consist of all zeroes except at the i_{th} position. Accordingly, this definition means, that a Boolean function is said to fulfill the avalanche criterion if and only if the average half of the output bits change, whenever one changes a single bit in the output bits. This implies that if a single input bit changes, then the output bits will change with 0.5 probability. According to

Adams, C., & Tavares, S the function of Hamming weight 2^{m-1} for all output, m-bits leads the S-box with the good avalanche. Because every vector f_j complements the input bit x_b according to alteration in the location from the position f_{j_i} and f_{j_k} for some positive integer j and k. If the vector f_j contains equal number ones and zeroes, then for all possible inputs with complementing the bit x_b yields the function y_i to be inverted 50%. Therefore, for all $f_1, f_2, \ldots, f_n$ with the property of hamming weight 2^{m-1} by inverting a bit x_b inverts on average half bits in $y_1, y_2, \ldots . y_n$. Since all the Boolean functions of the proposed S-boxes are complete, therefore the proposed S-boxes successfully satisfy SAC with an average probability approximately equal to 0.5 as can be seen in Tables 5a–5g.

Table 5: SAC Analysis of proposed S-boxes

(a) SAC analysis of S-box $\boldsymbol{S}_1$				(b) SAC analysis of S-box $\boldsymbol{S}_2$			
Result	Minimum	Maximum	Average	Result	Minimum	Maximum	Average
SAC	0.3750	0.6250	0.5069	SAC	0.4063	0.5938	0.5130
(c) SAC analysis of S-box $\boldsymbol{S}_3$				(d) SAC analysis of S-box $\boldsymbol{S}_4$			
Result	Minimum	Maximum	Average	Result	Minimum	Maximum	Average
SAC	0.437500	0.4965277	0.49652	SAC	0.37500	0.4904	0.49045
(e) SAC analysis of S-box $\boldsymbol{S}_4$				(f) SAC analysis of S-box $\boldsymbol{S}_5$			
Result	Minimum	Maximum	Average	Result	Minimum	Maximum	Average
SAC	0.43750	0.49998	0.49913	SAC	0.40625	0.49499	0.49479

5.4 Linear Approximation Probability

Linear approximation probability (LP) analysis is used to investigate the maximum value of the imbalance of the scheme. Let L_i and L_o be the input and the output mask respectively. According to the Mastui original definition of LP. The order of equal output bits selected by the mask L_o is equivalent to the equality of the input bits select by the mask L_i. Mathematically it can be written as follows:

$$LP = \max_{L_i, L_o \neq 0} \left| \frac{\{ i \in Z | i.L_i = S(i).L_o \}}{2^n} - \frac{1}{2} \right| \tag{12}$$

where the order of the set of in input value is 2^n. In Tables 6a–6f the values of maximum linear approximation probability of the S-boxes S_2, S_3 and S_4 are same that is equal to 0.187500. Similarly, the resultant linear approximation value of the S-box S_5 and S_6 are same equal to 0.2187500. The result of S_1 is equal to 0.25000 as shown in the tables. Since the probability results of all S-box are near zero therefor all S-boxes are secure against linear cryptanalysis.

Table 6: LP analysis of proposed S-boxes

(a). LP analysis of S-box $\boldsymbol{S}_1$				(b). LP analysis of S-box $\boldsymbol{S}_2$			
Result	Minimum	Maximum	Average	Result	Minimum	Maximum	Average
SAC	0	0.25000	0.04765	SAC	0	0.187500	0.04908

(Continued)

Table 6: Continued

(c). LP Analysis of S-box S_3				(d). LP Analysis of S-box S_4			
Result	Minimum	Maximum	Average	Result	Minimum	Maximum	Average
SAC	0	0.187500	0.04960	SAC	0	0.187500	0.04895
(e). LP Analysis of S-box S_5				(f). LP Analysis of S-box S_6			
Result	Minimum	Maximum	Average	Result	Minimum	Maximum	Average
SAC	0	0.218750	0.04882	SAC	0	0.218750	0.04895

The performance analyses demonstrate that the results of all analyses of the proposed S-boxes are quite better. According to Theorem 2, the APA transformations preserve the cryptographic properties of the S-box, so we used the APA transformation to increase the number of good quality S-boxes and robust their algebraic complexity. In the next section, we deployed the APA transformation in the Feistel network to enhanced the security of the DES algorithm.

6 Modified DES Algorithm

DES is a sixteen-round cryptosystem, each round is the combination of bits permutation, expansion of bits, substitution step and XOR operation. The practice of the bit permutation step is to rearrange the order of the data to aim to produce diffusion in the ciphered data. The usage of an exclusive XOR operation is to mix the round key with the plain data. The S-box is used to produce confusion in the ciphered data. In these operations the S-box is the only nonlinear component in the DES, thus the modification in any other operation of the algorithm would not make them less successful. Thus, in this study, we modified the DES algorithm by fitting a good quality 6-bit S-box in the FF and keep the other operation unchanged. The modified DES attains the following obligatory criteria.

i. Enhance the key space.
ii. Highly nonlinear output functions; the maximum distance from the linear functions.
iii. Successfully resist the linear and differential cryptanalysis.
iv. High nonlinearity is attained; degrees of the output bit functions are increased.
v. Efficient construction is easily implemented in hardware and software.

6.1 Generation of Key Dependent 6-Bits S-Boxes

The Key size of the modified DES is increased up to $12n$+56-bits. The first 56-bit of the key that is used to derive the sixteen round keys k_i. There is no change in deriving round keys. The last $12n$-bits are divided into $2n$ sub-blocks of 6-bits. Afterward, transforms the sub-blocks into the decimal form which is, of course, the elements of the Galois field $GF(2^6)$. The obtained elements are then used as parameters of APA transformation. For instance, let $a_1, a_2, \ldots, a_{2i}$ be the obtained elements. Then the APA transformation can be written as follows:

$$S(w) = (a_{2i-1}(\ldots a_7(a_5(a_3(a_1(w^{-1}) \oplus a_2) \oplus a_4) \oplus a_6) \oplus a_8 \ldots) \oplus a_{2i} \tag{13}$$

where w and a_i are the elements of the Galois field $GF(2^6)$ having a range between 0 and 63. Accordingly, for each different combination of $a_i \neq 0$, one can obtain different S-boxes of the same cryptographic properties and algebraic complexity. The purpose of the APA transformation is: First,

to increase the key space of the algorithm, increase the algebraic complexity of the S-box and generate a considerable number of key-dependent S-boxes having the same cryptographic properties. For the decryption, the same key generates the inverse of the S-box using the inverse of the transformation S given in Eq. (13).

6.2 Modified DES Feistel Network

The Feistel network was introduced by Horst Feistel. In general, it is a transformation of a function into a permutation, which is called F function. The FF is the nonlinear, reversible and key-dependent mapping, that maps an input string of data into the output string of data. The Feistel network has been widely used in many block ciphers such as in DES, GOST [13], Khufu and Khafre [14], RC5 [15], FEAL [17], Blowfish [18] and LOKI [19]. In this study, the Feistel network used in the DES is of our specific interest. The Feistel network plays a vital role in the security of DES. The input string of F-function in a round i is the right half output R_{i-1} of round $i-1$. The detailed procedure of the modified F-function is as: the modified F-function initially uses the E expansion and expands the 32-bit input data into 48-bit blocks. Afterward, the F-function carries out the XORed operation and mixed the 48-bit with the round key K_i. After the xor operation, the scheme divides the 48-bit block into eight 6-bit sub-blocks and substitutes each sub-block with the generated 6×6 S-box. The substitution method is: the first three most significant bits (MSB) selects the column of the S-box and the least significant three-bit (LSB) selects the row of the S-box. The output data are then again fed into eight different DES S-boxes. The detailed procedure of the modified Feistel network is demonstrated in Fig. 1.

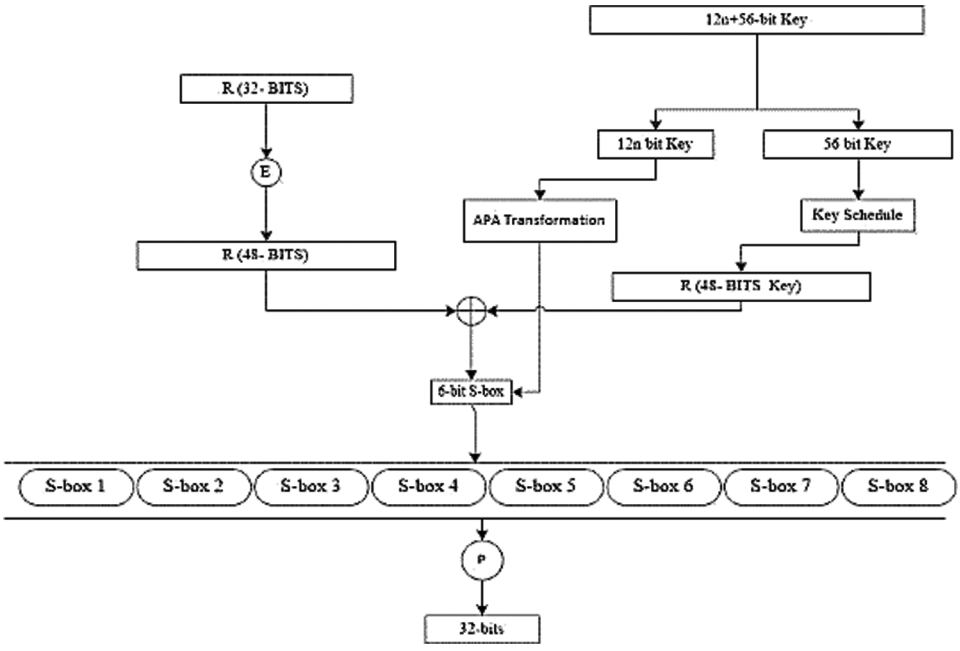

Figure 1: Flow Chart of the Modified F-function

Example: Let $I = 45$ be the input for the S-box S_1. The decimal representation of 45 is 101101_x. From the decimal representation, the MSB of the input I is $101_x = 5$ indicates the fifth row of the S-box S_1, the counts of the rows start from zero to 0. Similarly, the decimal representation of the input

LSB is again $101_x = 5$ indicates the fifth column of the S-box $\boldsymbol{S}_1$, the counts of the column start from zero 0. Thus, if the input I is substitute with the S-box given in Table 1, then the output of the S-box is $S_1(45) = 53$.

6.2.1 Key Compliment

The order of the key space of the DES algorithm is equal to 2^{52}. In that keyspace, half of the keys can be obtained by complimenting bitwise the other half keys. Since the DES cipher holds the following properties.

$$E(P,K) = C \quad \Rightarrow \quad \bar{C} = E(\bar{P},\bar{K}) \tag{14}$$

This property of the DES cipher makes the brute force attack simpler. The attacker has to check half possible keys to break the DES through a brute force attack. However, the S-boxes deployed in the modified DES are key-dependent, which does satisfy the following property.

$$S\,(P,k) = C \quad \not\Rightarrow \quad \bar{C} = S\left(\bar{P},\bar{k}\right) \tag{15}$$

Implies that

$$ME(P,K) = C \quad \not\Rightarrow \quad \bar{C} = ME(\bar{P},\bar{K}) \tag{16}$$

where ME denote the modified DES cipher, K denote the Modified DES key, S denote the substitution cipher of the scheme and k signify the subblock of the keys K that are used to generate the S-box S. Since the modified DES scheme does not hold the property given in Eq. (13). Hence, the attackers have to check all the keys in case of a brute force attack. We have examined the claim about the compliment property while using arbitrary key and plaintext on both DES and modified DES ciphers, the outcome is depicted in Table 7. From the table, it can be seen that the compliments of the DES cipher are equal to the ciphertext obtained as a result of using the key and plaintext compliment. However, the compliment of Modified DES ciphertext is not equal to the ciphertext given in the compliment row.

Table 7: Testing result

Data	Original	Compliment
Key	8, 9, 10, 11, 12, 13, 14, 15	247, 246, 245, 244, 243, 242, 241, 240
Plaintext	50, 54, 12, 43, 23, 54, 53, 55	205, 201, 243, 212, 232, 201, 202, 200
Ciphertext (DES)	126, 248, 50, 203, 126, 186, 50, 103	129, 7, 205, 52, 129, 69, 205, 152
Ciphertext (M DES)	133, 34, 22, 27, 234, 98, 194, 62	29, 16, 66, 205, 192, 5, 56, 74

6.2.2 The Brute Force Attack

A brute force attack is a classical attack, that is used to check all the possible keys until the correct key is found. In this era, symmetric ciphers with 100-bits key or less are susceptible to brute force attack. The DES algorithm uses 56-bit keys and therefore, it was proved to be insecure against brute

force attacks. The modified DES algorithm uses $12n$+56-bits. Accordingly, for $n \geq 5$, the algorithm will be able to resist the brute force attack. Since the modified DES algorithm is almost secure against linear and differential attacks, so the algorithm will be secure for $n = 1$, if all the round keys K_i are derived independently or by another complex method.

7 Conclusion

The DES algorithm was proved to be insecure against brute force attack, differential and linear cryptanalysis. The reason for breaking the algorithm was the small key space and weak S-box used in the substitution part of the algorithm. In this latter, we proposed an algorithm for the construction of 6×6 S-boxes based on Galois field $GF(2^6)$. The S-boxes are analyzed with different analyses and we found it secure against linear and differential attacks. Thus, we improved the DES algorithm by adding the construction method in the substitution part of the algorithm and strengthened the algorithm against brute force attack, linear and differential attacks.

Funding Statement: The authors extend their gratitude to the Deanship of Scientific Research at King Khalid University for funding this work through the research groups program under grant number R. G. P. 2/150/42.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] FIPS 46, "Data Encryption Standard," in *Federal Information Processing Standard*, Washington D. C., USA: National Bureau of Standards, U.S. Department of Commerce, 1977.

[2] W. Diffie and M. E. Hellman, "Special feature exhaustive cryptanalysis of the NBS data encryption standard," *Computer*, vol. 10, no. 6, pp. 74–84, 1977.

[3] M. E. Hellman, "A cryptanalytic time-memory trade-off," *IEEE Transactions on Information Theory*, vol. 26, no. 4, pp. 401–406, 1980.

[4] D. Chaum and J. H. Evertse, "Cryptanalysis of DES with a reduced number of rounds," in *Conf. on the Theory and Application of Cryptographic Techniques*, Berlin, Heidelberg, Springer, 1985.

[5] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *Journal of CRYPTOLOGY*, vol. 4, no. 1, pp. 3–72, 1991.

[6] T. Shah and D. Shah, "Construction of highly nonlinear S-boxes for degree 8 primitive irreducible polynomials over $\mathbb{Z}$ 2," *Multimedia Tools and Applications*, vol. 78, no. 2, pp. 1219–1234, 2019.

[7] F. A. Khan, J. Ahmed, J. S. Khan, J. Ahmad, M. A. Khan *et al.,* "A new technique for designing $8\times$ 8 substitution box for image encryption applications," in *2017 9th Computer Science and Electronic Engineering (CEEC)*, Colchester, UK: IEEE, pp. 7–12, 2017.

[8] M. A. Khan, J. Ahmad, Q. Javaid and N. A. Saqib, "An efficient and secure partial image encryption for wireless multimedia sensor networks using discrete wavelet transform, chaotic maps and substitution box," *Journal of Modern Optics*, vol. 64, no. 5, pp. 531–540, 2017.

[9] A. Qayyum, J. Ahmad, W. Boulila, S. Rubaiee, A. Arshad *et al.,* "Chaos-based confusion and diffusion of image pixels using dynamic substitution," *IEEE Access*, vol. 8, pp. 140876–140895, 2020.

[10] A. Arshad, S. Shaukat, A. Arshid, A. Eleyan, S. A. Shah *et al.,* "Chaos theory and its application: An essential framework for image encryption," *Chaos Theory and Applications*, vol. 2, no. 1, pp. 17–22, 2020.

[11] K. Nyberg, "Differentially uniform mappings for cryptography," in *Workshop on the Theory and Application of Cryptographic Techniques*, Springer, Berlin, Heidelberg, pp. 55–64, 1993.

[12] L. Cui and Y. Cao, "A new S-box structure named affine-power-affine," *International Journal of Innovative Computing, Information and Control*, vol. 3, no. 6, pp. 751–759, 2007.

[13] I. A. Zabotin, G. P. Glazkov and V. B. Isaeva, "Cryptographic protection for information processing systems: Cryptographic transformation algorithm." *Government Standard of the USSR*, Government Committee of the USSR for Standards, pp. 28147–28189, 1989.
[14] R. C. Merkle, "A fast software one-way hash function," *Journal of Cryptology*, vol. 3, no. 1, pp. 43–58, 1990.
[15] R. L. Rivest, "The RC5 encryption algorithm," in *Int. Workshop on Fast Software Encryption*, Springer, Berlin, Heidelberg, pp. 86–96, 1994.
[16] P. P. Mar and K. M. Latt, "New analysis methods on strict avalanche criterion of S-boxes," *World Academy of Science, Engineering and Technology*, vol. 2, no. 2, pp. 150–154, 2008.
[17] A. Shimizu and S. Miyaguchi, "Fast data encipherment algorithm FEAL," in *Workshop on the Theory and Application of Cryptographic Techniques*, Springer, Berlin, Heidelberg, pp. 267–278, 1988.
[18] B. Schneier, "Other block ciphers," in *Applied Cryptography*, Second Edition, New York, USA: John Wiley and Sons, pp. 319–325, 1996.
[19] L. P. Brown, M. Kwan, J. Pieprzyk and J. Seberry, "Improving resistance to differential cryptanalysis and the redesign of LOKI," in *Int. Conf. on the Theory and Application of Cryptology*, Springer, Berlin, Heidelberg, pp. 36–50, 1991.