IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 10, 2018, accepted December 24, 2018, date of publication January 18, 2019, date of current version February 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2893176

A Novel Cryptographic Substitution Box Design Using Gaussian Distribution

MUHAMMAD FAHAD KHAN[1], ADEEL AHMED[2,3], (Member, IEEE), AND KHALID SALEEM[3], (Member, IEEE)

[1]Department of Software Engineering, Foundation University, Islamabad 44000, Pakistan
[2]Department of Computer Science, National University of Modern Languages, Islamabad 44790, Pakistan
[3]Department of Computer Science, Quaid-i-Azam University, Islamabad 45320, Pakistan

Corresponding author: Muhammad Fahad Khan (fahad.khan@fui.edu.pk)

This work was supported by ORIC University.

ABSTRACT In this paper, a novel method is being proposed to construct a substitution box or Boolean function for block ciphers using Gaussian distribution and linear fractional transform. The substitution box is constructed by employing a linear fractional transform based on Box–Muller transform, polarization decision, and central limit algorithm. The cryptographic strength of the proposed S-boxes is evaluated with standardized tests such as linear approximation probability, unified averaged changed intensity, bit independent criterion, histogram analysis, nonlinearity score, strict avalanche criterion, and differential approximation probability. The results show that the proposed substitution box achieves better cryptographic strength as compared with the state-of-the-art techniques.

INDEX TERMS Substitution permutation networks, block cipher, cryptographic confusion, S-box, random number generation, Gaussian distribution, linear fractional transform.

I. INTRODUCTION

Substitution box is a nonlinear primitive of block cipher which is basically a set of permutations mapping of m-bits inputs to n-bits output. The n-bits output can be viewed as a boolean function such as F :$\mathrm{F}_2^{\mathrm{n}} \longrightarrow \mathrm{F}_2^{\mathrm{m}} F : F_2^n \longrightarrow F_2^m$ [1], [2]. Generally, s-box is used to hide the relationship between the key and cipher-text and usually it is a single non-linear transformation component which performs confusion of bits as described in literature [2], [3], [41]. There are two adopted structures for the block cipher, substitution permutation networks (SPN) and feistel kind of networks [1]. The SPN structure is widely used in advanced encryption standard (AES) which employs bijective S-boxes in order to make the encryption algorithm invertible. The feistel network is used in data encryption standard (DES) which is not restricted to bijective S-boxes and can use non-bijective mapping as well. For block ciphers such as AES and DES, liner and differential attacks are considered as powerful attacks. These attacks can only be resisted if S-box attains properties such a high non-linearity and low differential uniformity. Generally, for the evaluation of the encryption system, the two basic design criteria suggested by Ratiner et al. are diffusion and confusion [2]. In the diffusion method, if we change a single bit of plain text, then numerous bits of encrypted bits should be changed whereas in confusion method each bit of the encrypted text depends on numerous bits of the key [4]. The most basic AES based S-box properties are defined in literature [3], [5]:

- Nonlinearity: No substitution box is linear affine method of its input.
- Every row is permutation: Frequency of zeros and ones must be same.
- Avalanche property: At least 2 bits of the output must be changed, if single bit of input is changed.
- If one input bit is fixed, S-box minimizes the difference of zeros and ones on the output.

Maximum achievable high nonlinearity score for even n-bits based substitution boxes are called vectorial bent functions and can only exist for $m \leq {}^{n}\!/_{2}$ [1]. The output distributions of all derivatives of the vectorial bent functions are uniform but it is not the balanced S-box [3]. Maximum, high non-linearity score is 116 till now. The efficiency of the block cipher symmetric encryption schemes is totally depend upon the designing method of the S-boxes. General methods for the construction of S-boxes are based on chaotic maps, power polynomial, DNA sequences, TDERC sequences, galois field, machine learning, inversion mapping and pseudo-random number generator [6], [9]. Over the past decade, researchers have shown great interest in studying the behavior of chaotic system for S-boxes construction and in existing literature indicates chaotic maps have been extensively used.

TABLE 1. Comparison of Gaussian and Chaos theory techniques.

NIST TESTS	Ginger bread man Map	Henon Map	Lorenz Map	Logistic Map	Chen System	Box–Muller Transform	Polarization Decision	Central Limit Algorithm
Frequency Test	P	P	P	--	P	P	P	P
Test For Frequency Within A Block	--	--	--	--	--	P	P	P
Runs Test	P	--	P	--	P	P	P	P
The Longest Run Of Ones In A Block	--	--	--	P	-	p	P	P
Random Binary Matrix Rank Test	--	P	--	--	--	P	--	--
Non-Overlapping Template Matching	--	--	--	--	--	P	P	P
Maurer's Universal Statistical	P	P	P	--	P	--	--	--
Cumulative Sum Test	P	P	-	--	--	P	P	P
Discrete Fourier Transform test	--	--	--	P	P	P	P	P
Rank Test	--	--	P	--	--	P	P	P

These systems exhibits many favorable characteristics for encryption such as continuous broad-band power spectrum, strong randomness, state periodicity, non-convergence and extreme sensitivity to initial conditions. Cryptographic strength of the chaotic cipher depends upon the chaotic system and encryption scheme. Chaotic systems are divided into two classes; one dimensional and multi-dimensional. Depending upon the number of dependent or independent parameters each type of the chaotic classes has its own demerits. In literature [10], the most used chaotic maps for S-boxes design are Ginger bread man, Henon, Lorenz, Logistic, and Chen maps. Many researchers have also pointed weakness in behavior of chaotic systems like non uniform distribution of data, discontinuity in chaotic sequences, short quantity of randomness, finite precision effect, and computational complexity [10], [30], [37], [55], [56]. Gaussian distribution is also available to generate pseudo random numbers [11], [14]. Gaussian distribution base random number generators can generate random numbers of sufficient quality to meet the demands of a particular simulation environment [11]. There are four categories of Gaussian random number generators such as cumulative density function (CDF), inversion transformation, rejection and recursive methods. CDF methods simply invert the CDF to produce random numbers from the given distribution. Transformation methods transform the uniform numbers into a Gaussian distribution. Rejection is similar to transformation but have an additional step of conditionally rejecting some of the transformed values. Recursive methods utilize linear combination of previously generated Gaussian numbers to produce new random numbers [11]. Mostly used Gaussian distribution algorithms for random number generation are; Box-Muller transform, polarization decision and central limit [11], [13]. Assessment of chaotic and gaussian base random numbers are shown in Table 1. For comparison, we selected mostly used chaotic maps and gaussian distribution base random number generation techniques to evaluate the quality of pseudo randomness using NIST randomness suit.

We can see that Gaussian distribution based pseudo random numbers show better results as compared to chaos based pseudo random numbers. Although chaos provides efficient cryptographic properties but still need further improvements because cryptanalysis techniques are improving day by day. To cope the pace of cryptanalysis advancement, recently published studies used hybrid chaotic S-box construction approaches. Because single chaotic map based techniques have failed to achieve higher standards. So, for the proposed design we selected Box-Muller transform, Polarization decision, and Central limit algorithms of Gaussian distribution instead of chaotic maps [15], [16]. The mathematical description of each method is given below:

The Box-Muller transform method is the exact transformation method which generates pair of random numbers from two uniform numbers [13], [15], [16]. It is based on the property of a two-dimensional Cartesian system in which X and Y coordinates are described by two independent and normally distributed random variables such as

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{x^2}{2\delta^2}} \tag{1}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{y^2}{2\delta^2}} \tag{2}$$

The generated pair random numbers r and θ from these two X, Y coordinates are independent and described as

$$f_R(r) = \int_0^{2\pi} \frac{r}{2\pi\delta^2} e^{-\frac{r^2}{2\delta^2}} d\theta, \quad 0 \leq \mathrm{r} \leq \infty \tag{3}$$

$$f_\Theta(\theta) = \int_0^{\infty} \frac{r}{2\pi\delta^2} e^{-\frac{r^2}{2\delta^2}} dr, \quad 0 \leq \theta \leq 2\pi \tag{4}$$

R obeys the Rayleigh distribution and Θ obeys the uniform distribution, so their joint probability density is $f_{R\Theta}(r, \theta) = f_R(r) \times f_\Theta(\theta)$ which is also statistically independent. So the corresponding distribution for R and Θ can be described as

$$F_R(r) = \int_0^r \frac{r'}{\delta^2} e^{-\frac{r^2}{2\delta^2}} dr' \tag{5}$$

$$F_\Theta(\theta) = \int_0^\theta \frac{1}{2\pi} d\theta' = \frac{\theta}{2\pi} \tag{6}$$

Both $F_R(r)$ and $F_\Theta(\theta)$ are in closed form and hence Gaussian random variables can be generated by the inverse transformation method.

The polarization decision is another precise approach to obtain the two-dimensional Gaussian distribution. It is related to Box-Muller transform method but is superior to it. It considers pair of random number having same property as required in Box-Muller method and then the probability density function is given as

$$F = \int_{-\infty}^{+\infty} f_X(x)\,dx = \int_{-\infty}^{+\infty} f_Y(y)dy \tag{7}$$

Thus the square of F transformed to polar coordinates, is defined by equation 8.

$$\begin{aligned} F^2 &= \frac{1}{2\pi\delta^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2\delta^2}}\, d(x)\, d(y) \\ &= \frac{1}{2\pi\delta^2}\int_0^{2\pi}\int_0^{+\infty} re^{-\frac{r^2}{2\delta^2}} dr d\theta \end{aligned} \tag{8}$$

This is similar to Box-Muller transform method, the transformation to polar coordinates makes the θ uniformly distributed from 0 to 2π. The normalized distribution function of radial distance r is:

$$P(r < a) = \int_0^a re^{-\frac{r^2}{2}} dr \tag{9}$$

Here, a uniform random number U is also used having values in the interval of [0, 1]. A new point is generated by multiplying that point by radial distance r : (rcos (2πU), rsin (2πU)) and their inverse transformation gives the two standard normal variables.

Central limit is the efficient method for generating Gaussian random numbers, since it simply samples sufficient amount of identical and independent uniform distributions. The arithmetic means of their distribution will have normal distribution. More formally it can be described by assuming "n" number of independent and identically distributed uniform numbers $U_i \sim U(0, 1)$ then sum of U_i is described as

$$S = \sum_{i=1}^{n} U_i \tag{10}$$

The cumulative distribution function of S can be approximated as:

$$F_S(s) = \Phi(\frac{s - n\mu}{\sqrt{n\sigma^2}}) \tag{11}$$

The mean μ and variance σ are given by $^1/_2$ and $^1/_{12}$ respectively, where Φ represents the cumulative distribution function of Gaussian distribution. If choose variable z such as

$$z = \frac{s - \frac{n}{2}}{\frac{1}{12}\sqrt{n}} \tag{12}$$

The distribution function of z is Gaussian distribution:

$$f_Z(z) = \frac{n}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{13}$$

After normalization, a standard Gaussian distribution is obtained.

In this paper we use the Gaussian distribution methods to generate random numbers and then use these numbers in our design. Rest of the paper is organized as follows; section II present our contribution, section III describes proposed methodology, section IV, explain the results and evaluation of proposed method and section V gives the conclusion.

II. CONTRIBUTION

In this paper, our major contributions are:

a- New method is proposed for the construction of S-box design using Gaussian distribution.

b- The results show that generated S-box resistivity against attacks in terms of maximum non-linearity is similar to Advanced Encryption Standard (AES) substitution box.

c- The substitution box based encryption technique passes all cryptanalysis tests, substitution box security evaluation criteria, differential analysis tests and histogram analysis.

III. PROPOSED DESIGN METHODLOGY

We proposed a novel design methodology to construct substitution box using linear fractional transformation that is

$$f(z) = \frac{az + b}{cz + d} \tag{14}$$

where a ∈ Box-Muller transform, b ∈ polarization decision algorithm, c ∈ central limit algorithm. In addition, the conditions of $cz = -d$ and $ad - bc \neq 0$ are avoided. Proposed design steps are given below:

Step 1: For variable "a", "b", "c": generate sequence from three Gaussian distribution algorithms (Box-Muller transform, polarization decision algorithm, and central limit algorithm). Figure-1a is the proposed design to generate a, b and c.

Step 2: Remove floating points from these sequences and get three sequences of random numbers.

Step 3: Convert these three sequences of random numbers into their respective binary representation.

Step 4: Select LSB from each binary sequence and generate parity bits stream.

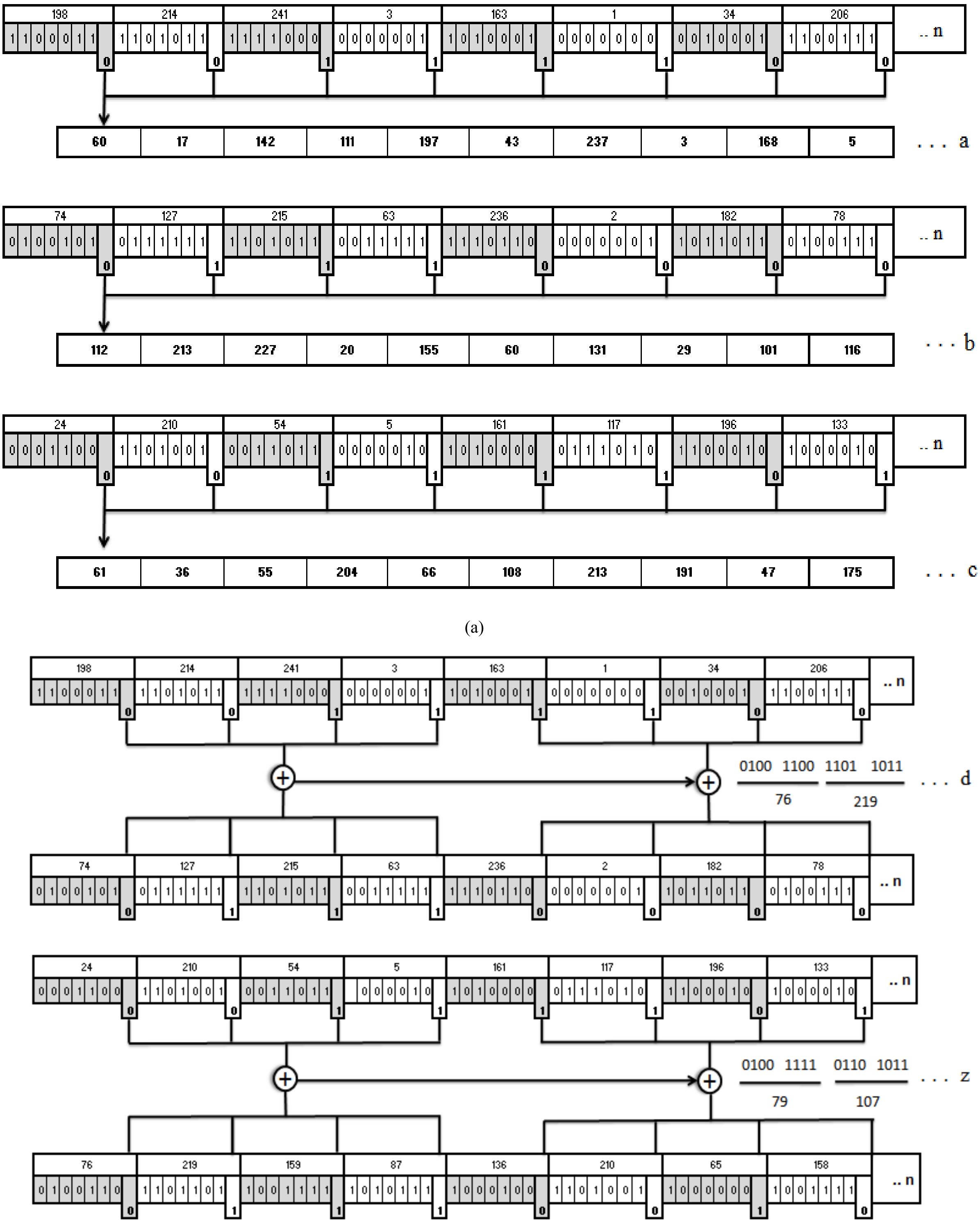
Step 5: Combine 8 parity bits and convert it into their respective decimal numbers.

Step 6: For variable "d": (a) perform the bitwise XOR on parity bits generated from Box-Muller transform method and (b) polarization decision method.

Step 7: Convert the resultant sequence into their respective binary representation.

Step 8: Combine 8 bits and convert it into their respective decimal numbers.

Step 9: For variable "z": (c) perform the bitwise XOR on parity bits generated from central limit method and



(a)

(b)

FIGURE 1. **(a) Proposed design to generate a, b and c. (b) Proposed design to generate d and z.**

sequences generated for "d" in step 8. Figure-1b is the proposed design to generate d and z.

Step 10: Convert the resultant sequence into their respective binary.

Step 11: Combine 8 parity bits and convert it into their respective decimal numbers.

Step 12: Perform linear fractional transformation which is defined in equation (14).

]

TABLE 2. Proposed S-box.

26	179	99	22	131	250	43	52	85	164	171	197	156	6	36	160
89	118	147	81	18	78	86	61	53	213	127	146	0	230	216	150
108	39	113	153	101	48	13	205	201	54	165	166	151	116	66	64
211	109	173	47	70	62	236	115	92	238	249	217	204	111	21	189
4	181	17	75	185	34	228	30	252	155	207	218	41	186	220	229
59	210	196	134	202	187	23	91	183	224	163	240	14	74	2	19
128	234	32	138	175	37	49	190	100	192	35	180	184	159	206	11
140	29	231	28	167	7	170	199	112	110	226	222	144	38	174	232
40	161	225	135	107	209	114	88	102	126	119	95	141	84	12	145
105	133	178	60	71	24	25	58	57	69	154	221	87	77	1	215
33	50	188	79	31	44	237	208	122	169	16	132	94	214	200	63
55	254	139	83	176	96	162	152	129	80	194	235	246	193	191	67
233	239	227	104	247	143	203	223	253	51	9	255	248	130	251	123
241	73	177	106	65	46	124	117	136	82	243	76	242	10	149	5
3	98	182	148	172	158	42	219	212	120	8	45	244	72	168	125
198	137	142	90	68	195	56	97	245	103	27	157	121	93	20	15

The numbers obtained as a result of f (z) are chunk by chunk asses with random sizes and then transformed in substitution box which is shown in Table 2.

A. RESULTS AND EVALUATION

This section is about results analysis and evaluation of proposed S-box via different evaluation criteria such as, s-box security evaluation criteria, cryptography analysis, differential analysis and histogram analysis. Our results show that proposed method achieves 112 maximum non-linearity which is similar to ASE algorithm. Similarly, other statistical analysis also proved that our encryption scheme, based on proposed S-box has also high resistivity against attacks. The analysis of S-box is described in following steps:

B. S-BOXES SECURITY EVALUATION CRITERIA

1) NON-LINEARITY

The distance between the function and the set of all affine functions is called nonlinearity. It can be represented as the number of bits that must be changed in the truth table of a Boolean to achieve the closest affine function [18]. The ability of the cryptographic function resistance against the linear attacks can be denoted by its non-linearity score [17]. So, larger value of that property is required. The proposed substitution box maximum, minimum and average scores are 112,110,111, respectively.

We can see that our maximum achieved non-linearly score is greater and equal to 30 recently published research papers as shown in Table 3.

TABLE 3. Nonlinearity of various S-boxes.

Substitution box	Max. Non linearity	Substitution box	Max. Non linearity
Daemen [57], 2002	112	Jakimoski[6],2001	98
Ye [27], 2018	112	Silva[42],2018	106
Zhang[1], 2014	112	Sarfraz[43],2016	106
Gangadari[28], 2016	110	Solami[44],2018	108
Wang[9], 2010	106	Lambić[45],2017	108
Zhang[36], 2018	110	Tian[46],2018	106
Ahmad[37], 2016	110	Lambić[47],2014	112
Islam[38], 2017	108	Ahmad[51],2015	110
Çavuşoğlu[39], 2017	106	Lambic[8],2018	108
Belazi[53], 2017	108	Belazi[53],2017	110
Çavuşoğlu[39], 2017	110	Belazi[54],2017	112
Özkaynak[40], 2018	108	Zhang[49],2014	110
Liu [50], 2015	108	Liu [48], 2014	106

TABLE 4. Bit independent criterion.

----	112	108	112	112	112	108	110
112	----	108	110	110	110	112	110
108	108	----	110	108	110	110	110
112	110	110	----	110	112	112	110
112	110	108	110	----	112	110	110
112	110	110	112	112	----	112	110
108	112	110	112	110	112	----	110
110	110	110	110	110	110	110	----

2) BIT INDEPENDENT CRITERION (BIC)

Another standard for evaluation of S-box is output bits independence criterion which is a desirable property for every cryptography system. BIC property can be explained as all the avalanche variables should be pair-wise independent for a given set of avalanche vectors, generated by the complementing of a single plaintext bit.

The BIC property results for the proposed S-box is shown in tables 4 and 5, showing that the proposed S-box satisfies

TABLE 5. BIC dependent MATRIX.

----	.51758	.49219	.48437	.48242	.52344	.52929	.52344
.51758	----	.51172	.51562	.48242	.48242	.52734	.48633
.49219	.51172	----	.50781	.49219	.51172	.50195	.51953
.48437	.51562	.50781	----	.50976	.49609	.49609	.49414
.48242	.48242	.49219	.50976	----	.50000	.50195	.52148
.52344	.48242	.51172	.49609	.50000	----	.50000	.51758
.52929	.52734	.50195	.49609	.50195	.50000	----	.50390
.52344	.48633	.51953	.49414	.52148	.51758	.50390	----

BIC property close to the best possible value. Here minimum BIC score is 108.

3) STRICT AVALANCHE CRITERIA (SAC)

The confusion ability of the S-box is analyzed on the basis of SAC results. The SAC results of S-box will be satisfied if one bit change in the input results in a change in half of the output bits. Formally it can be defined as function $f : F_2^n \longrightarrow F_2$ satisfies SAC if $f(x) \oplus f(x \oplus \alpha)$ is balanced for all α whose weights are 1. The SAC result of the proposed S-box is given in table 6.

TABLE 6. Strict avalanche criteria.

.50000	.48437	.50000	.45312	.54687	.45312	.46875	.53125
.46875	.50000	.46875	.54687	.46875	.46872	.54687	.48437
.50000	.45312	.43750	.53125	.46875	.48437	.48437	.50000
.51562	.51562	.50000	.54687	.51562	.50000	.56250	.53125
.51562	.48437	.54687	.53125	.51562	.54687	.51562	.51562
.46875	.53125	.53125	.50000	.54687	.50000	.51562	.43750
.56250	.53125	.51562	.53125	.51562	.50000	.43750	.53125
.50000	.50000	.48437	.51562	.46875	.54687	.46875	.48437

Basically, SAC depicts information that once a single unit of eight length byte of plaintext converted from 0 to 1, the altering likelihood of each binary unit in the output is 0.5. The proposed S-box achieves minimum, maximum and average values for SAC i.e. 0.437500, 0.562500, and 0.503662 respectively, and variance is 0.032075.

C. DIFFERENTIAL ANALYSIS

Differential attack is a plaintext chosen attack in which attacker analyzes the results that come back to the known cipher text. Number of changing pixel rate (NPCR) and the Unified averaged changed intensity (UACI) are the two most common evaluation indicators which have ability to test resistance against differential attack. These tests are discussed as follows:

1) NPCR ANALYSIS

Number of changing pixel rates represented as:

$$NPCR = \frac{\sum_{i,j} D(i,j)}{N \times M} \times 100\% \tag{15}$$

TABLE 7. Comparative NPCR and UACI analysis of proposed method with AES.

Algorithms	NPCR	UACI
Proposed	99.60	33.09
Wang [25], 2018	99.59	33.45
Sathish [20], 2011	98.47	32.21
Huang [21], 2009	99.42	24.94
Huang [22], 2013	99.54	28.27
Fouda [23], 2014	99.60	33.42
Loukhaoukha [19], 2012	99.58	28.62
Guo [26], 2018	99.60	33.46
Hussain [24], 2018	99.30	33.40

TABLE 8. Comparison of NPCR and UACI of the proposed scheme for lena test image.

Images	Loc.	NPCR		UACI	
		Proposed	AES	Proposed	AES
Cameraman	First	99.62	99.61	30.57	33.54
	Mid	99.64	99.62	37.40	33.53
	Last	99.61	99.59	3425	33.53
Lena	First	99.60	99.61	33.40	33.40
	Mid	99.61	99.66	31.95	33.32
	Last	99.60	99.61	33.94	33.52
Baboon	First	99.75	99.624	30.40	33.45
	Mid	99.93	99.61	28.95	33.46
	Last	99.89	99.62	31.30	33.53

2) UACI ANALYSIS

Unified Averaged Changed Intensity is represented as:

$$UACI = \frac{1}{N \times M} \times \left[\sum\nolimits_{i,j} \frac{|C_1(i,j) - C_2(i,j)|}{255} \right] \times 100\%$$

$$f(x) = \begin{cases} 0, & \mathit{if}\ \mathrm{C_1(i,j) = C_2(i,j)}, \\ 1, & \mathrm{if\ C_1(i,j) \neq C_2(i,j)}, \end{cases} \tag{16}$$

Table 7 and table 8 shows the comparative analysis of proposed method with ASE.

D. CRYPTANALYSIS

1) LINEAR APPROXIMATION PROBABILITY (LP)

The LP is the maximum value of the imbalance of an event. The parity of the input bits selected by the mask Γx is equal to the parity of the output bits selected by the mask Γy. The smaller the LP, the stronger the ability of the S-box resistance against linear cryptanalysis attacks. For LP calculation, we used the following definition as given in [17] and [18]

$$LP_f = max_{\Gamma x, \Gamma y \neq 0} \left| \frac{\{x \in X \,|\, x.\Gamma x = \mathrm{S}(x) \cdot \Gamma y\}}{2^n} - \frac{1}{2} \right| \tag{17}$$

Γx and Γy are the input and output mask respectively, where X is the set of all possible inputs; and 2^n is the number of its elements. The maximum value of the LP for our proposed S-box is 0.078125.

2) DIFFERENTIAL APPROXIMATION PROBABILITY (DP)

In order to resist the differential type of cryptanalysis attacks, the S-box should have differential uniformity. The smaller

TABLE 9. **Differential approximation probability.**

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.0	4.0

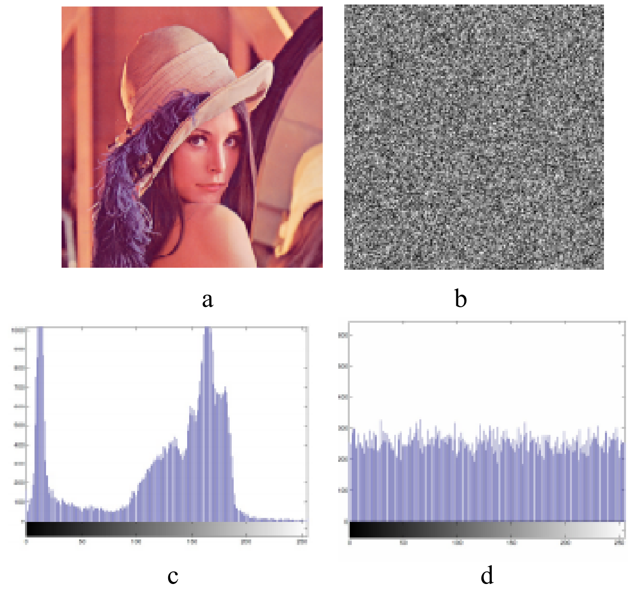


FIGURE 2. **Lena plain image, 2b: Encrypted image of Lena, 2c: Lena plain image histogram 2d: Encrypted Lena image histogram.**

the DP, the stronger is the ability of the S-box, resisting against differential cryptanalysis attacks. DP can be a sign of the XOR sharing of the initial seed and output of the Boolean function. In DP, the input differential Δx_i should uniquely map to an output differential Δy_i, therefore ensuring a uniform mapping probability for each i. For DP calculation, we used the following definition of DP as given in [17] and [18] and the DP result of our proposed S-box is shown in table 9.

$$DP\,(\Delta x \longrightarrow \Delta y) = \left[\frac{\#\{x \in X \mid (\mathrm{S(x)} \oplus S\,(x \oplus \Delta x) = \Delta y\}}{2^n} \right] \tag{18}$$

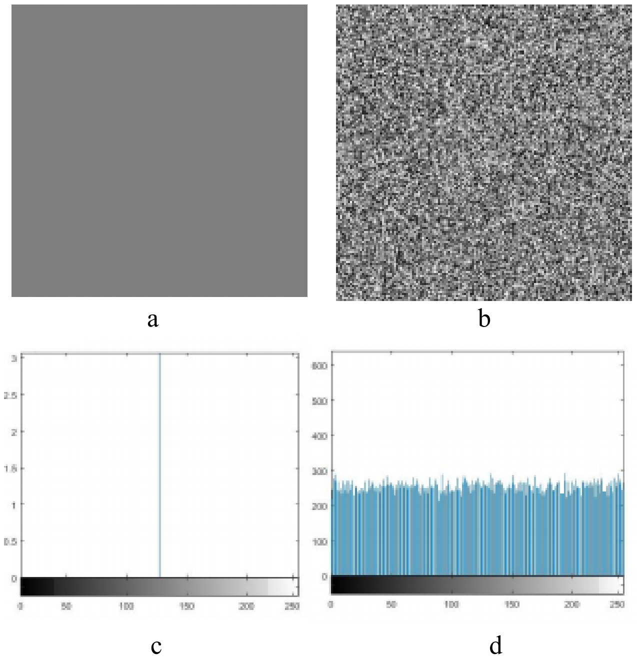


FIGURE 3. **a) Plain one gray scale image, b) One gray scale image in encrypted form, 3c: Plain one gray scale image histogram, 3d: Encrypted one gray scale image histogram.**

E. HISTOGRAM ANALYSIS

Performance evaluation of encrypted image using histogram analysis is an important parameter. Efficient image encryption algorithm encrypts the plain image into image that contains random pixels.

Histogram analysis illustrates how pixels are distributed after encryption process. Plain image of Lena is shown in figure 2a and histogram of Lena image in shown in figure 2c.

Encrypted Lena image from our proposed s-box is show in figure 2b. Figure 2b is totally distorted and gives no clue of original image. Our results are further confirmed in figure 2d, in which uniform distributed histogram of cipher image is shown.

For further analysis, we also test our proposed method on gray image as shown in figure 3a, having 125 gray values, and its histogram is shown in figure 3c. Encrypted gray image from our proposed S-box is shown in figure 3b. In 3b we can see that pixels of encrypted image are entirely random. Also histogram of encrypted gray image is again uniformly distributed.

IV. CONCLUSION

Random numbers are the important part of encryption methods. Recently authors used chaos theory and Gaussian distribution methods to generate random numbers. These random numbers can be used to construct S-boxes. To select the most appropriate random number generator for S-box, we compare the quality of the random numbers generated by these two approaches using NIST randomness tests. After comparison, we selected Gaussian distribution methods. Chaos based S-box provides efficient cryptographic properties but still need further improvements because cryptanalysis techniques are improving day by day. To cope with the pace of cryptanalysis advancement, different studies now use hybrid chaotic based S-box construction approaches, because single chaotic based techniques sometime fail to achieve randomness standards. In this paper, we used Gaussian distribution methods and their pseudorandom numbers to construct the S-box. Our analysis proves that constructed S-box pose good cryptographic properties. Proposed methodology can generate N size of S-boxes or boolean functions with random quality. In future, we will also utilize proposed methodology to generate diffusion.

REFERENCES

[1] W. Zhang and E. Pasalic, "Highly nonlinear balanced S-boxes with good differential properties," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7970–7979, Dec. 2014.

[2] M. Ratiner, "The method of S-box construction," *J. Discrete Math. Sci. Cryptogr.*, vol. 8, no. 2, pp. 203–215, 2005.

[3] K. Mohamed, M. N. M. Pauzi, F. H. H. M. Ali, S. Ariffin, and N. H. N. Zulkipli, "Study of S-box properties in block cipher," in *Proc. Int. Conf. Comput., Commun., Control Technol. (I4CT)*, Sep. 2014, pp. 362–366.

[4] J. Szczepanski, J. M. Amigo, T. Michalek, and L. Kocarev, "Cryptographically secure substitutions based on the approximation of mixing maps," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 2, pp. 443–453, Feb. 2005.

[5] L. Kocarev, "Chaos-based cryptography: A brief overview," *IEEE Circuits Syst. Mag.*, vol. 1, no. 3, pp. 6–21, Mar. 2001.

[6] G. Jakimoski and L. Kocarev, "Chaos and cryptography: Block encryption ciphers based on chaotic maps," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 2, pp. 163–169, Feb. 2001.

[7] T. T. K. Hue, T. M. Hoang, and D. Tran, "Chaos-based S-box for lightweight block cipher," in *Proc. IEEE 5th Int. Conf. Commun. Electron. (ICCE)*, Jul. 2014, pp. 572–577.

[8] D. Lambić, "S-box design method based on improved one-dimensional discrete chaotic map," *J. Inf. Telecommun.*, vol. 2, no. 2, pp. 181–191, 2018.

[9] Y. Wang, L. Yang, M. Li, and S. Song, "A method for designing S-box based on chaotic neural network," in *Proc. 6th Int. Conf. Natural Comput. (ICNC)*, vol. 2, Aug. 2010, pp. 1033–1037.

[10] A. Ullah, S. S. Jamal, and T. Shah, "A novel scheme for image encryption using substitution box and chaotic system," *Nonlinear Dyn.*, vol. 91, no. 1, pp. 359–370, 2018.

[11] D. B. Thomas, W. Luk, P. H. W. Leong, and J. D. Villasenor, "Gaussian random number generators," *ACM Comput. Surv.*, vol. 39, no. 4, p. 11, 2007.

[12] D. B. Thomas, "FPGA Gaussian random number generators with guaranteed statistical accuracy," in *Proc. IEEE 22nd Annu. Int. Symp. Field-Program. Custom Comput. Mach. (FCCM)*, May 2014, pp. 149–156.

[13] Y. Hu, Y. Wu, Y. Chen, and G. C. Wan. (2018). "Gaussian random number generator based on FPGA." [Online]. Available: https://arxiv.org/abs/1802.07368

[14] G. Zhang, P. H. W. Leong, D.-U. Lee, P. H. W. Leong, J. D. Villasenor, R. C. C. Cheung, and W. Luk, "Ziggurat-based hardware Gaussian random number generator," in *Proc. Int. Conf. Field Program. Logic Appl.*, 2005, pp. 275–280.

[15] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, "Quantum random number generation," *npj Quantum Inf.*, vol. 2, p. 16021, Jun. 2016.

[16] T. Shinzato, "Box muller method," Hitotsubashi Univ., Kunitachi, Japan, Tech. Rep. A30-23-2007, 2007.

[17] A. Altaleb, M. S. Saeed, I. Hussain, and M. Aslam, "An algorithm for the construction of substitution box for block ciphers based on projective general linear group," *AIP Adv.*, vol. 7, no. 3, p. 035116, 2017.

[18] I. Hussain, T. Shah, H. Mahmood, M. A. Gondal, and U. Y. Bhatti, "Some analysis of S-box based on residue of prime number," *Proc. Pak Acad. Sci.*, vol. 48, no. 2, pp. 111–115, 2011.

[19] K. Loukhaoukha, J.-Y. Chouinard, and A. Berdai, "A secure image encryption algorithm based on Rubik's cube principle," *J. Elect. Comput. Eng.*, vol. 2012, Jan. 2012, Art. no. 7.

[20] G. A. Sathishkumar, K. Bhoopathy, and R. Sriraam, "Image encryption based on diffusion and multiple chaotic maps," *Int. J. Netw. Secur. Appl.*, vol. 3, pp. 181–194, Mar. 2011.

[21] C. K. Huang and H.-H. Nien, "Multi chaotic systems based pixel shuffle for image encryption," *Opt. Commun.*, vol. 282, pp. 2123–2127, Jun. 2009.

[22] C. K. Huang, C.-W. Liao, S. L. Hsu, and Y. C. Jeng, "Implementation of gray image encryption with pixel shuffling and gray-level encryption by single chaotic system," *Telecommun. Syst.*, vol. 52, no. 2, pp. 563–571, 2013.

[23] J. S. Fouda, A. Eyebe, J. Y. Effa, S. L. Sabat, and M. Ali, "A fast chaotic block cipher for image encryption," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 19, no. 3, pp. 578–588, 2014.

[24] I. Hussain, A. Anees, A. H. AlKhaldi, A. Algarni, and M. Aslam, "Construction of chaotic quantum magnets and matrix Lorenz systems S-boxes and their applications," *Chin. J. Phys.*, vol. 56, no. 4, pp. 1609–1621, 2018.

[25] X. Wang, X. Zhu, and Y. Zhang, "An image encryption algorithm based on Josephus traversing and mixed chaotic map," *IEEE Access*, vol. 6, pp. 23733–23746, 2018.

[26] J.-M. Guo, D. Riyono, and H. Prasetyo, "Improved beta chaotic image encryption for multiple secret sharing," *IEEE Access*, vol. 6, pp. 46297–46321 , 2018.

[27] Y. Ye, N. Wu, X. Zhang, L. Dong, and F. Zhou, "An optimized design for compact masked AES S-box based on composite field and common subexpression elimination algorithm," *J. Circuits, Syst. Comput.*, vol. 27, no. 11, p. 1850171, 2018.

[28] B. R. Gangadari and S. R. Ahamed., "Design of cryptographically secure AES like S-box using second-order reversible cellular automata for wireless body area network applications," *Healthcare Technol. Lett.*, vol. 3, no. 3, pp. 177–183, 2016.

[29] H. Jiang *et al.*, "A novel true random number generator based on a stochastic diffusive memristor," *Nature Commun.*, vol. 8, no. 1, p. 882, 2017.

[30] J. Gayathri and S. Subashini, "A survey on security and efficiency issues in chaotic image encryption," *Int. J. Inf. Comput. Secur.*, vol. 8, no. 4, pp. 347–381, 2016.

[31] G. Zhao, G. Chen, J. Fang, and G. Xu, "Block cipher design: Generalized single-use-algorithm based on chaos," *Tsinghua Sci. Technol.*, vol. 16, no. 2, pp. 194–206, 2011.

[32] N. Hadj-Said, B. Belmeki, and A. Belgoraf, "Chaotic behavior for the secrete key of cryptographic system," *Chaos, Solitons Fractals*, vol. 23, no. 5, pp. 1549–1552, 2005.

[33] G. Jakimoski and L. Kocarev, "Analysis of some recently proposed chaos-based encryption algorithms," *Phys. Lett. A*, vol. 291, no. 6, pp. 381–384, 2001.

[34] N. K. Pareek, V. Patidar, and K. K. Sud, "Discrete chaotic cryptography using external key," *Phys. Lett. A*, vol. 309, nos. 1–2, pp. 75–82, 2003.

[35] Y. Lu, L. Li, H. Zhang, and Y. Yang, "An extended chaotic maps-based three-party password-authenticated key agreement with user anonymity," *PLoS ONE*, vol. 11, no. 4, p. e0153870, 2016.

[36] T. Zhang, C. L. P. Chen, L. Chen, X. Xu, and B. Hu, "Design of highly nonlinear substitution boxes based on I-Ching operators," *IEEE Trans. Cybern.*, vol. 48, no. 12, pp. 3349–3358, Dec. 2018.

[37] M. Ahmad, N. Mittal, P. Garg, and M. M. Khan, "Efficient cryptographic substitution box design using travelling salesman problem and chaos," Perspect. Sci., vol. 8, pp. 465–468, Sep. 2016.

[38] F. Islam and G. Liu, "Designing S-box based on 4D-4wing hyperchaotic system," *3D Res.*, vol. 8, no. 1, p. 9, 2017.

[39] Ü. Çavuşoğlu, A. Zengin, I. Pehlivan, and S. Kaçar, "A novel approach for strong S-Box generation algorithm design based on chaotic scaled Zhongtang system," *Nonlinear Dyn.*, vol. 87, no. 2, pp. 1081–1094, 2017.

[40] F. Özkaynak, "Construction of robust substitution boxes based on chaotic systems," *Neural Comput. Appl.*, pp. 1–10, Nov. 2018.

[41] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, 1949.

[42] V. M. Silva-García, R. Flores-Carapia, C. Rentería-Márquez, B. Luna-Benoso, and M. Aldape-Pérez, "Substitution box generation using Chaos: An image encryption application," *Appl. Math. Comput.*, vol. 332, pp. 123–135, Sep. 2018.

[43] M. Sarfraz I. Hussain, and F. Ali, "Construction of S-box based on Mobius transformation and increasing its confusion creating ability through invertible function," *Int. J. Comput. Sci. Inf. Secur.*, vol. 14, no. 2, p. 187, 2016.

[44] E. Al Solami, M. Ahmad, C. Volos, M. Doja, and M. Beg, "A new hyperchaotic system-based design for efficient bijective substitution-boxes," *Entropy*, vol. 20, no. 7, p. 525, 2018.

[45] D. Lambić, "A novel method of S-box design based on discrete chaotic map," *Nonlinear Dyn.*, vol. 87, no. 4, pp. 2407–2413, 2017.

[46] Y. Tian and L. Zhimao, "Chaotic S-box: Six-dimensional fractional Lorenz–Duffing chaotic system and O-shaped path scrambling," *Nonlinear Dyn.*, vol. 94, no. 3, pp. 2115–2126, 2018.

[47] D. Lambić, "A novel method of S-box design based on chaotic map and composition method," *Chaos, Solitons Fractals*, vol. 58, pp. 16–21, Jan. 2014.

[48] H. Liu, A. Kadir, and Y. Niu, "Chaos-based color image block encryption scheme using S-box," *AEU-Int. J. Electron. Commun.*, vol. 68, no. 7, pp. 676–686, 2014.

[49] X. Zhang, Z. Zhao, and J. Wang, "Chaotic image encryption based on circular substitution box and key stream buffer," *Signal Process., Image Commun.*, vol. 29, no. 8, pp. 902–913, 2014.

[50] G. Liu, W. Yang, W. Liu, and Y. Dai, "Designing S-boxes based on 3-D four-wing autonomous chaotic system," *Nonlinear Dyn.*, vol. 82, no. 4, pp. 1867–1877, 2015.

[51] M. Ahmad, D. Bhatia, and Y. Hassan., "A novel ant colony optimization based scheme for substitution box design," *Procedia Comput. Sci.*, vol. 57, pp. 572–580, Jan. 2015.

[52] T. Farah, R. Rhouma, and S. Belghith, "A novel method for designing S-box based on chaotic map and teaching–learning-based optimization," *Nonlinear Dyn.*, vol. 88, no. 2, pp. 1059–1074, 2017.

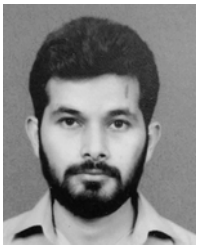
[53] A. Belazi, A. A. A. El-Latif, A.-V. Diaconu, R. Rhouma, and S. Belghith, "Chaos-based partial image encryption scheme based on linear fractional and lifting wavelet transforms," *Opt. Lasers Eng.*, vol. 88, pp. 37–50, Jan. 2017.

[54] A. Belazi, M. Khan, A. A. A. El-Latif, and S. Belghith, "Efficient cryptosystem approaches: S-boxes and permutation-substitution-based encryption," *Nonlinear Dyn.*, vol. 87, no. 1, pp. 337–361, 2017.

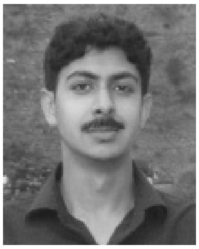
[55] C. Li, D. Arroyo, and K.-T. Lo, "Breaking a chaotic cryptographic scheme based on composition maps," *Int. J. Bifurcation Chaos*, vol. 20, no. 8, pp. 2561–2568, 2010.

[56] C. Li, S. Li, and K.-T. Lo, "Breaking a modified substitution–diffusion image cipher based on chaotic standard and logistic maps," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 16, no. 2, pp. 837–843, 2011.

[57] J. Daemen and V. Rijmen, *The Design of Rijndael: AES—The Advanced Encryption Standard*. Berlin, Germany: Springer-Verlag, 2002.



MUHAMMAD FAHAD KHAN is currently an Assistant Professor with Foundation University and a Ph.D. Scholar with the Department of Computer Science, Quaid-i-Azam University, Islamabad. He has authored more than 30 research papers. His research interests include steganography, cryptography, and multimedia communication.



ADEEL AHMED received the M.Phil. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan, in 2011, where he is currently a Ph.D. Scholar with the Department of Computer Science. He was granted a travel award in HUGO 15th Human Genome Meeting, Dubai, United Arab Emirates, in 2011. He is currently a Faculty Member of Computer Science, National University of Modern Languages, Islamabad. He has authored more than ten research papers. His research interests include large scale complex schema matching and integration using machine learning, recommendation systems, social network analysis, and bioinformatics with a focus on simulation techniques.

KHALID SALEEM received the M.Sc. degree from Quaid-i-Azam University, Pakistan, in 1994, and the M.Phil. and Ph.D. degrees from the University of Montpellier 2, France, in 2005 and 2008, respectively, all in computer science. He was with the software industry from last 20 years. He is currently an Assistant Professor with Quaid-i-Azam University. He is the President of PAK-France Alumni Network. He has authored more than 20 research papers and various scientific reports and book chapters. His research interests include data mining, schema matching and integration, database systems, bioinformatics, distributed systems, and data warehousing.