Algebraic Construction and Cryptographic Properties of Rijndael Substitution Box

Shristi Deva Sinha and Chaman Prakash Arya
Institute for Systems Studies and Analyses, Delhi – 110 054, India
E-mail: shristideva_issa@yahoo.com

ABSTRACT

Rijndael algorithm was selected as the advanced encryption standard in 2001 after five year long security evaluation; it is well proven in terms of its strength and efficiency. The substitution box is the backbone of the cipher and its strength lies in the simplicity of its algebraic construction. The present paper is a study of the construction of Rijndael Substitution box and the effect of varying the design components on its cryptographic properties.

Keywords: Advanced encryption standard, Rijndael, Rijndael algorithm, substitution box, cryptographic properties

1. INTRODUCTION

The advanced encryption standard (AES) developed by Joan Daemen and Vincent Rijmen, Rijndael\(^1,2\) was selected by National Institute of Standards and Technology (NIST) as in 2001. It is a symmetric block cipher-based on the Shannon substitution-permutation network. AES has long been an area of interest for the researchers due to its:

- Well proven security: A 5-year evaluation procedure by NIST and it is designed to be resistant to linear, differential and mount attacks.
- Efficiency: The speed of encryption and decryption of AES is the fastest compared to any other cipher of similar strength.
- Design simplicity: The cipher has a simple and elegant structure that can easily be split into its components.

Most of the earlier works relating to AES are linked to its performance evaluation and straightforward implementation\(^3-7\) including the various pipelined architecture. Some simplification\(^8,9\) in the AES algorithm had been attempted. However, these simplifications lead to vulnerabilities\(^10\) in the algorithm. A great amount of work has also been done in fast pipelined implementation\(^11-15\) of the algorithm. Rijmen proposal of AES S-box implementation based on the composite fields\(^16\) was a significant step to compact AES. Some work in optimum construction\(^17\) of these composite fields has been done. Some study on the replacement\(^18\) of the design parameters of the Rijndael algorithm has been done. It was suggested that this leads to creation of new ciphers equivalent in strength to the original. However, certain properties\(^19\) of substitution box (S-box) has been identified, which are profoundly affected by the changes in design components. Recent works relating to AES S-box include the optimised implementation of the S-box using residues of prime numbers\(^20\), a lightweight mix columns implementation for AES\(^21\) and a proposal of a new algorithm to construct secure keys for AES\(^22\) is published.

This paper focuses on the study of the algebraic construction of the S-box of the AES algorithm, which is the main strength of the cipher. The effect of the change in the design components of the S-box on its cryptographic properties has been analysed. It provides an insight to the AES S-box construction to generate a conceptual framework for all future customizations of the algorithm targeted at the S-box design level.

2. BASIC STRUCTURE OF ADVANCED ENCRYPTION STANDARD

Full description of the AES algorithm can be obtained in FIPS\(^2\) 197. The input and the output for AES are each bit sequences containing 128-bits. However, AES allows cipher keys of all 128-bits, 192-bits, or 256-bits lengths. The input 128 bits are arranged in a 4×4 matrix, termed the ‘state’ and all byte operations are performed in the Galois field GF(2\(^{8}\)). The cipher is specified in terms of repetitions of processing steps that are applied to make up rounds of keyed transformations between the input plain-text and the final output cipher-text. A set of reverse rounds are applied to transform cipher-text back into the original plain-text using the same encryption key. The encryption of data is done in number of rounds:

- a. Initial round: AddRoundKey
- b. Rounds: SubBytes, ShiftRows, MixColumns, and AddRoundKey
- c. Final round: SubBytes, ShiftRows, and AddRoundKey.

\textit{Addroundkey} : Addroundkey is a XOR of the key with the array.
\textit{ShiftRows} : ShiftRows cyclically shifts the elements of the
$i^{th}$ row of the state $C$, elements to the right, where, $C_i$ are fixed constants $C_i = 0, 1, 2$ and $3$.

**MixColumns** : The columns of the state are considered as polynomials over $GF(2^n)$ and multiplied modulo $x^4+1$ by $03.x^3+x^2+02$ to give a new column array.

**SubBytes** : SubBytes is a nonlinear byte substitution, operating on each of the state bytes independently.

In all these operations, the SubBytes deserve a special mention as this step involves the S-box. The S-box is the backbone of the cipher; it provides nonlinearity in the encryption process and plays an important role in key scheduling.

### 3. ALGEBRAIC PRELIMINARIES AND S-BOX CONSTRUCTION

For the study of algebraic construction of the S-box a theorem is stated here without proof.

**Theorem 1**

Let $p$ be a non-zero element of a principle ideal domain $R$ then, $R(p)$ will be a field if and only if, $p$ is irreducible.23

According to this theorem, for a prime $p$ Galios field $GF(p^n)$ is constructed by using a generating polynomial $m(x)$ of degree $n$ taking

$$GF(p^n) = \frac{GF(p)[x]}{m(x)}$$

In the AES algorithm, the irreducible polynomial $x^8 + x^4 + x^3 + x + 1$ is used to generate the underlying field $GF(2^8)$. All bytes $b$ in Rijndael are interpreted as elements of this field represented by a polynomial $a_i + a_2x + a_3x^2 + a_4x^3 + a_5x^4 + a_6x^5 + a_7x^6 + a_8x^7$ where, each bit $a_i \in GF(2)$ and $b \in GF(2^8)$. In this field, addition $\oplus$ and multiplication $\odot$ are defined by the XOR operation and polynomial multiplication modulo the generating polynomial respectively.

An S-box is a transformation $\sigma : GF(p^n) \rightarrow GF(p^n)$, $p$ is a prime

In AES the S-box $\sigma : GF(2^8) \rightarrow GF(2^8)$ is constructed by substituting each element with its inverse and applying a suitable affine transformation $\sigma : X \rightarrow AX^{-1} + b$ where, $A \in GL_8(2)$, the general linear group of degree 8 over $GF(2)$ and $b \in GF(2^8)$.

Both of these $A$ and $b$ are fixed in AES:

$$A = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1
\end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix}
1 \\
0 \\
0 \\
1 \\
0 \\
1 \\
0 \\
0
\end{bmatrix}$$

The motivation for this S-box design is to be resistant to differential, linear cryptanalysis and interpolation attacks. The core design is a simple transformation $x \rightarrow x^{-1}$ in $GF(2^n)$, this mapping has a simple algebraic expression. However, the simplicity itself makes it vulnerable to attacks like the interpolation attack. Therefore, it is combined with a suitable affine transformation: $x \rightarrow Ax^{-1} + b$. The affine mapping is so chosen that, it has a very simple description, but a complicated algebraic expression. If combined with the ‘inverse’ mapping, it can be seen as modular polynomial multiplication followed by addition $b(x) = (x^7 + x^6 + x^2 + x + a(x)) (x^7 + x^6 + x^5 + x^4 + 1) \mod (x^8 + 1)$. The purpose of the constant translation vector $b$ is to ensure that $\exists$ no fixed and conjugate fixed points (i.e. $\exists$ no $x \in GF(2^n)$ such that $\sigma(x) = x$ or $\sigma(x) = \overline{x}$) in the S-box.

### 4. CHARACTERISTICS OF A GOOD S-BOX AND BIAS PARAMETERS

To study the characteristics of a good S-box $\sigma : GF(p^n) \rightarrow GF(p^n)$, it is realised as vectorial Boolean function $\sigma(x) = (\sigma_1(x), \sigma_2(x), ..., \sigma_n(x))$ where, each $\sigma_i$ is a Boolean function of the Boolean variables $x_1, x_2, ..., x_n$.

Characteristics of the good S-box are, it has to:

- be balanced
- satisfy propagation criterion
- satisfy correlation immunity criterion
- have input/output bit-to-bit entropy ($H = 1$).
- its nonlinearity ($N$) has to be $\geq 120$.

Each of these criteria’s are explained in the following subsection. However, most of these are satisfied only for an ideal case and not in practical level. So, biases from each of these criteria are derived.

#### 4.1 Balancedness Property

This property states that each of the Boolean functions of the S-box should be balanced, i.e., the number of ones and zeros in the truth table of the Boolean function must be equal.

#### 4.2 Propagation Criterion

A Boolean function is said to satisfy propagation criterion of degree $k$ and order $m$, if any function obtained by keeping $m$ input bits fixed $f(x)$ changes with probability half, whenever $i$ ($1 \leq i \leq k$) bits of $x$ are complemented.

Mathematically, by fixing $m$ number of bits $C_{m}2^m$ set of functions $g$ are obtained from $f$, let it be denoted by $F$. Let $\alpha \in GF(2^n)$: $W(\alpha) \in [1, k]$ then a function is said to satisfy the propagation criterion of degree $k$ and order $m$, if for each $g \in F$, $g(x) \oplus g(x \oplus \alpha)$ is balanced.

The propagation criterion is the measure of randomness of the differences in output pairs to the input pairs. This is a very important criterion as the bias of the distribution of the differences of the output pairs and the input pairs is utilised in the differential cryptanalysis of the conventional ciphers.

**Definition**

Propagation criterion bias of a Boolean function of degree $k$ and order $m$ is defined by:\

$$\text{ Defined as }$$
A total of 30 irreducible polynomials of degree 8 are computed and the parameter is called nonlinearity. The effectiveness of these functions is measured by a parameter where values of these parameters for the AES S-box are: H = 0.9887, PCB(1,0) = 16, PCB(1,1) = 20, CIB(1) = 16 and nonlinearity N = 112. The values of these bias parameters are used to analyze the effect of changes in the design components of the AES S-box on its cryptographic properties. Different possible variations on the S-box components and their effects have been discussed in the next section.

5. ANALYSIS AND RESULTS
The S-box is constructed by the transformation:

\[ x = Ax^{-1} + b, \quad x \in GF(2^8), \]
\[ A \in GL_4(2) \quad \text{and} \quad b \in GF(2^8). \]

All the variations in its construction without altering the simple algebraic expression are looked into and their effects in the bias parameter values are analyzed. One of the major changes that can be brought about without altering the algebraic expression is by changing the underlying field to isomorphic fields. Another option is to change the affine matrix A and third is changing the vector b.

5.1 Change in the Underlying Field to Isomorphic Fields
Isomorphic fields to the underlying field can be generated by using different irreducible polynomials of the same degree. Number of irreducible polynomials of degree 8 over GF(p) is given by:

\[ \frac{1}{n} \sum_{d \mid n} \mu(d)p^{n/d}, \]
where \( \mu(n) \) is the Mobius function.

Thus, \( \exists \) a total of 30 irreducible polynomials of degree 8 over GF(2) including the one originally used. The irreducible polynomials are constructed by generating the polynomials and testing their irreducibility using the following theorem:

Theorem 2: Rabin’s Test for irreducibility
A polynomial \( C \in GF(p)[x] \) of degree d is irreducible if and only if \( X^d \equiv x \mod C \)

The irreducible polynomials and the values of the bias parameters of the respective S-boxes constructed on isomorphic field generated by them are shown in Table 1.

It can be observed from Table 1 that values of the bias parameters can be enhanced on changing to isomorphic fields as in the case of the irreducible polynomial \( x^8 + x^6 + x^5 + x + 1 \).
Table 1. Generating polynomials and the corresponding parameter values of the S-boxes

<table>
<thead>
<tr>
<th>Generating Polynomials</th>
<th>H</th>
<th>PCB (1,0)</th>
<th>PCB (1,1)</th>
<th>CIB(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x^8 + x^4 + x^3 + x + 1 ) (AES)</td>
<td>0.9887</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^5 + x^4 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^6 + x^5 + x^4 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^4 + x^3 + x^2 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^4 + x^2 + 1 )</td>
<td>0.9887</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^6 + x^5 + x^3 + x^2 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^5 + x^3 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^5 + x^4 + x^3 + x^2 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + x^2 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^5 + x^4 + x^3 + x^2 + 1 )</td>
<td>0.9937</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^6 + x^5 + x^3 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + x^2 + 1 )</td>
<td>0.9937</td>
<td>12</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^5 + x^4 + x^3 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^6 + x^5 + x^3 + x^2 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + x^2 + x + 1 )</td>
<td>0.9937</td>
<td>16</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^3 + x^2 + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^6 + x^5 + x + 1 )</td>
<td>0.9914</td>
<td>12</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^5 + x^2 + x + 1 )</td>
<td>0.9887</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x^3 + x^2 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^2 + x + 1 )</td>
<td>0.9887</td>
<td>12</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>( x^8 + x^7 + x^6 + x + 1 )</td>
<td>0.9914</td>
<td>16</td>
<td>20</td>
<td>14</td>
</tr>
</tbody>
</table>

5.2 Change in the Affine Matrix

Another variation in the S-box design component can be brought about by changing the affine matrix \( A \). The affine matrix \( A \in GL_8(2) \), the general linear group of degree 8 over GF(2) and the order of this group is:

\[
\prod_{k=0}^{7}(2^8 - 2^k) = 5.3481 \times 10^{18}.
\]

Hence, the number of matrix \( A \) available to be used for varying the S-box is numerous. So, to analyse the effect of such change on the bias parameter values instead of performing an exhaustive search over this group, a random search has been done. For this, a square binary matrix of size 8 is randomly generated. The matrix is discarded, if found to be singular and another matrix is generated again. This non-singular randomly generated matrix is used in the construction of S-box and the bias parameters are computed. This process is implemented in MatLab and is repeated 500 times. Few randomly generated affine matrices obtained on 500 such random searches are:
Some of the observed bias parameter values obtained due to the change in the affine matrix in the random search are shown in Table 2.

### Table 2. Values of the bias parameters obtained for different affine matrix

<table>
<thead>
<tr>
<th>H</th>
<th>PCB(1,0)</th>
<th>PCB(1,1)</th>
<th>CIB(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.989</td>
<td>16</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>0.991</td>
<td>16</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>0.991</td>
<td>14</td>
<td>20</td>
<td>12</td>
</tr>
</tbody>
</table>

#### 5.3 Change in the Translation Vector

The translation vector $b \in GF(2^8)$ and $O(GF(2^8)) = 256$ therefore, $\exists$ total 256 different vector $b$ (including the one used in the AES) that can be used in the S-box construction. However, it is observed that change in the translation vectors has no impact on the values of the bias parameters. Only the number of fixed and conjugate fixed points of the S-box varies with it. It was observed that the nonlinearity ($N=112$) of the AES S-box was not affected by any of the above mentioned changes.

#### 6. CONCLUSIONS

Rijndael algorithm accepted as the AES is a well proven and efficient cipher. The S-box forms a backbone of this cipher and is designed with a very simple algebraic expression. This paper studies the construction of its S-box and explores the possible design variations without altering its simple algebraic expression. The effect of these changes on the cryptographic properties of the S-box has also been analysed. The main aim of this study is to provide an insight to the AES S-box construction to generate a conceptual framework for all possible customization on the cipher targeted at the S-box design level. It is observed that by changing the underlying field into the isomorphic fields improves the properties to some level. A similar effect is observed for the change in the affine matrix. However, change in the translation vector shows no effect in the cryptographic properties of the S-box.

#### ACKNOWLEDGEMENTS

The authors would like to thank Shri H.V. Srinivasa Rao Director, Institute for Systems Studies and Analyses, DRDO for his encouragement and support in this work. The authors would also like to thank Shri P.K. Bhatnagar and Shri Yogesh Chandra for their kind suggestions and guidance.

#### REFERENCES


Contributors

Mr Shristi Deva Sinha received his MSc (Mathematics) from University of North Bengal, Darjeeling. Currently, he is working as a Scientist in the Institute for Systems Studies & Analyses (ISSA), DRDO, Delhi. He has been involved in the analysis of naval weapon systems / procedures and cryptography.

Mr Chaman Prakash Arya received his MSc (Mathematics) from University of Delhi, Delhi. He is presently pursuing PhD in General Topology from University of Delhi. Currently, he is working as a Scientist in the ISSA, DRDO, Delhi. He has been involved in the area of OLI evaluation of weapons and cryptography.