Journal of Informatics and Mathematical Sciences
Vol. 10, Nos. 1 & 2, pp. 339–350, 2018
ISSN 0975-5748 (online); 0974-875X (print)
Published by RGN Publications
http://dx.doi.org/10.26713/jims.v10i1-2.662

http://www.rgnpublications.com

Research Article

An Algorithm for Improving Algebraic Degree of S-Box Coordinate Boolean Functions Based on Affine Equivalence Transformation

Hoang Duc Tho[1], Nguyen Truong Thang[2], Nguyen Thi Thu Nga[2], Pham Quoc Hoang[1]

[1] *Academy of Cryptographic Techniques, Ha Noi, Viet Nam*
[2] *Viet Nam Academy of Science and Technology - Institute of Information Technology, Ha Noi, Viet Nam*
***Corresponding author:** hdth@bcy.gov.vn

Abstract. The Substitution box (S-box) plays an important role in a block cipher as it is the only nonlinear part of the cipher in most cases. S-box S can be considered as a vectorial Boolean function consisting of m individual Boolean functions $f_1, f_2, \ldots, f_m$, where $f_i : \mathrm{GF}(2^n) \to \mathrm{GF}(2)$ and $f_i(x) = y_i$ for $i = 1, 2, \ldots, m$. These functions are called coordinate Boolean functions of the S-box S. To avoid various attacks on the ciphers and for efficient software implementation, the coordinate Boolean functions of S-boxes are required to satisfy a lot of properties, for instance being a permutation defined on the fields with even degrees, with a high nonlinearity, a low differential uniformity and a high algebraic degree, etc. However, it seems very difficult to find an S-box with the coordinate Boolean functions to satisfy all the criteria. The S-box with low algebraic degree of the coordinate Boolean functions is vulnerable to many attacks such as linear and differential cryptanalysis, for instance higher-order differential attacks, algebraic attacks or cube attacks. In this paper we propose an algorithm for improving algebraic degree of the S-box coordinate Boolean functions while not affecting its other important properties. The algorithm is based on affine equivalence transformation of the S-boxes.

Keywords. S-boxes; Affine equivalence; Algebraic degree

MSC. 11T71; 14G50; 68P25

Received: July 5, 2017 **Accepted:** January 10, 2018

1. Introduction

In [19], C. Shannon defined the confusion and diffusion which should exist in an encryption system. In block ciphers, S-box and P-box are two important components of a secure block cipher identified by C. Shannon. The basic purpose of an S-box is to produce confusion between the ciphertext and the secret key and P-box is responsible for diffusion. Basically, confusion is required so that the ciphertext is related to both the plaintext and secret key, in a complex way. Since the S-box plays an important role in a block cipher as it is the only nonlinear part of the cipher in most cases. To avoid various attacks on the ciphers and for efficient software implementation, S-boxes are required to satisfy a lot of properties, for instance being a permutation defined on the fields with even degrees, with a high algebraic degree, a low differential uniformity and a high nonlinearity, etc., [20, 12, 14]. However, it seems very difficult to find an S-box to satisfy all the criteria.

Actually, the S-box with the coordinate Boolean functions of low algebraic degree is vulnerable to many attacks, for instance higher-order differential attacks, algebraic attacks or cube attacks [4, 5]. In this paper we propose an algorithm that allows improving algebraic degree of the S-box coordinate Boolean functions while not affecting its other properties in order to increase ability to resist the attacks mentioned above. This proposed algorithm is based on affine equivalence transformation of S-boxes [14, 1].

The rest of the paper is organized as follows: In Section 2, some fundamental definitions are given as a refresher, to help better understand the research results given later in the article, and the main cryptographic properties of S-box are discussed. In Section 3, we present and discuss our algorithm. Finally, Section 4 summarizes the paper.

2. Definitions and preliminaries

In this section we will briefly recall some of the basic definitions and properties of Boolean functions. For a comprehensive survey on Boolean functions we refer to [6, 7, 14].

S-box: Let the S-box of an n-binary input into m-binary output mapping is denoted by S. Then $S : \mathrm{GF}(2^n) \to \mathrm{GF}(2^m)$ and to each $x = (x_1, x_2, \ldots, x_n) \in \mathrm{GF}(2^n)$ some $y = (y_1, y_2, \ldots, y_m) \in \mathrm{GF}(2^m)$ is assigned by $S(x) = y$, where $\mathrm{GF}(2) = \{0, 1\}$ is the 1-dimensional Boolean space. Clearly, S can be considered as a vectorial Boolean function consisting of m individual Boolean functions $f_1, f_2, \ldots, f_m$, where $f_i : \mathrm{GF}(2^n) \to \mathrm{GF}(2)$ and $f_i(x) = y_i$ for $i = 1, 2, \ldots, m$. These functions are called coordinate Boolean functions of the S-box S and it is well known that most of the desirable cryptographic properties of S can be defined in terms of their linear combinations. The S-box coordinate Boolean functions and all their linear combinations are referred as the S-box component Boolean functions.

Although the original Data Encryption Standard [16] used S-boxes mapping a six-bit input to a four-bit output, most modern cipher designs use only bijectiven $\times nS$-boxes. In particular, we note that the current Advanced Encryption Standard (AES) [15], and most block ciphers designed for environments where the AES is too resource-intensive (e.g. PRESENT [2]), use only bijectiven $\times nS$-boxes. For this reason, we will focus primarily on functions of this sort.

Algebraic normal form of Boolean function: A Boolean function can be represented by a truth table, which is the binary output vector of the function containing 2^n elements. We obtain the polarity truth table when instead of $f(x)$, the signed function $\widehat{f}(x) = (-1)^{f(x)}$ is considered. Another way of representing a Boolean function is bymeans of its *algebraic normal form* (ANF).

A Boolean function f can be represented uniquely in the form of a polynomial with 2^n coefficients whose degree in each variable is at most 1 (*Zhegalkin polynomial* or *Algebraic Normal Form-ANF*) [11, 3]:

$$f(x) = a_0 \oplus \sum_{i=1}^{n} a_i x_i \oplus \sum_{1 \le i < j \le n} a_{ij} x_i x_j \oplus \ldots \oplus a_{1\ldots n} x_1 \ldots x_n, \tag{2.1}$$

where all the 2^n coefficients $a_i, a_{ij}, \ldots, a_{1\ldots n} \in \mathrm{GF}(2)$ and $x_i \in \{0,1\}$, for all $i = 1, \ldots n$, and $\sum$, $\oplus$ denote the modulo 2 summation.

Each coefficients of the ANF $a_0, a_1, \ldots, a_i, a_{ij}, \ldots, a_{1\ldots n}$ corresponds to term in the ANF form.

Let us represent the polynomial (2.1) through a vector p, in which all the coefficients $a_i, a_{ij}, \ldots, a_{1\ldots n}$ are replaced by coefficients $a_0, a_1, \ldots, a_{2^n-1}$, respectively, where $a_i \in \mathrm{GF}(2)$, for all $i = 0 \ldots 2^n - 1$.

The algebraic degree of an n-variable Boolean function $f(x)$, denoted by $\deg(f)$, is the number of variables of the largest product term of the function's ANF having a non-zero coefficient.

For example, with p in the form of the coefficient vector $p = (1,0,1,0,0,1,1,0)$, then the corresponding polynomial is $f(x) = 1 \oplus x_1 \oplus x_0 x_3 \oplus x_1 x_2$ the algebraic degree of $f(x)$ is 2.

There exist various generalizations of the concept of algebraic degree to the case of the S-box. The most common such definition [8, 12] is as follows:

The algebraic degree ($\deg(S)$) of an S-box S to be the minimum of the degrees of all non-zero linear combinations of its component functions:

$$\deg(S) = \min\{\deg(c_1 f_1 \oplus c_2 f_2 \oplus \ldots \oplus c_m f_m)\}$$

where $c = (c_1, c_2, \ldots, c_m) \in \mathrm{GF}(2^m) \setminus \{0\}$.

For example, the S-box used in PRESENT [2] is an 4-bit to 4-bit S-box. The action of this S-box in hexadecimal notation is given by Table 1.

Table 1. The PRESENT S-box

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
S(x)	C	5	6	B	9	0	A	D	3	E	F	8	4	7	1	2

The coordinate Boolean functions in the ANF form and their algebraic degree are presented in Table 2 [10].

Table 2. The coordinate Boolean functions of PRESENT S-box

Boolean functions in ANF	Algebraic degree
$p_1 = (0,1,0,0,1,0,1,0,1,0,0,0,0,0,0,0)$ $f_1(x) = x_0 \oplus x_2 \oplus x_1x_2 \oplus x_3$	2
$p_2 = (0,0,1,0,0,0,0,1,1,0,1,1,1,1,0,0)$ $f_2(x) = x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3
$p_2 = (1,0,0,1,1,0,0,0,1,1,1,1,0,1,0,0)$ $f_3(x) = 1 \oplus x_0x_1 \oplus x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
$p_4 = (1,1,1,0,1,0,1,1,1,0,0,1,0,1,0,0)$ $f_4(x) = 1 \oplus x_0 \oplus x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3

We computed all their linear combinations from these four original coordinate Boolean functions of the S-box and determined algebraic degree of this S-box that is 2.

To convert a Boolean function as a truth table into the ANF, we can multiply the Hardamad matrix with vector of Boolean function values. This method is described in detail in [17, 18, 8].

The Walsh-Hadamard transform (WHT) of an n-variable Boolean function $\widehat{f}(x)$, denoted by $\widehat{F}_f(w)$, is defined by:

$$\widehat{F}_f(w) = \sum_{x \in \mathrm{GF}(2^n)} \widehat{f}(x)(-1)^{\langle w,x \rangle} = \sum_{x \in \mathrm{GF}(2^n)} (-1)^{f(x) \oplus \langle w,x \rangle} = \sum_{x \in \mathrm{GF}(2^n)} \widehat{f}(x)\widehat{l}_w(x),$$

where $\widehat{l}_w(x)$ is the signed function of the linear function $l_w(x) = \langle w,x \rangle$.

Thus, for all $w \in \mathrm{GF}(2^n)$, $\widehat{F}_f(w) \in [-2^n, 2^n]$. $\widehat{F}_f(w)$ is called a spectral Walsh coefficient and the real-valued vector of all 2^n Walsh coefficients is referred to as the WHT Spectrum.

The autocorrelation transform (ACT) of $\widehat{f}(x)$, denoted by $\widehat{r}_f(\alpha)$, taken with respect to a vector $\alpha \in \mathrm{GF}(2^n)$ is defined by:

$$\widehat{r}_f(\alpha) = \sum_{x \in \mathrm{GF}(2^n)} (-1)^{f(x) \oplus f(x \oplus \alpha)} = \sum_{x \in \mathrm{GF}(2^n)} \widehat{f}(x)\widehat{f}(x \oplus \alpha).$$

Thus, for all $\alpha \in \mathrm{GF}(2^n)$, $\widehat{r}_f(\alpha) \in [-2^n, 2^n]$ and $\widehat{r}_f(0) = 2^n$. The $\widehat{r}_f(\alpha)$ is called a spectral autocorrelation coefficient and the real-valued vector of all 2^n autocorrelation coefficient representing the ACT of function is referred to as the ACT Spectrum.

Nonlinearity: Let $c = (c_1, c_2, \ldots, c_n)$ be a nonzero element in $\mathrm{GF}(2^n)$. Let $c \cdot S = c_1f_1 \oplus c_2f_2 \oplus \ldots \oplus c_nf_n$ be a linear combination of the coordinate Boolean functions $f_1, f_2, \ldots, f_n$ of S. The nonlinearity (NL) for an S-box is defined as:

$$NL(S) = \min_{c \in \mathrm{GF}(2^n), c \neq 0} NL(c \cdot S)$$

The NL of S is the Hamming distance between the set of all non-constraint linear combinations of component functions of S and the set of all affine functions over GF(2).

Differential Uniformity: The differential uniformity of an $n \times nS$-box S, denoted by δ, is defined as the largest value present in its difference distribution table (DDT) not counting the first entry in the first row. That is,

$$\delta = \max_{a \in \mathrm{GF}(2^n) \setminus \{0\}} \max_{b \in \mathrm{GF}(2^n)} |\{x \in \mathrm{GF}(2^n) | S(x) \oplus S(x \oplus a) = b\}|$$

Then, S is said to be differentially δ-uniform.

Fixed point: An element $x \in \mathrm{GF}(2^n)$ is a fixed point of S-box $S : \mathrm{GF}(2^n) \to \mathrm{GF}(2^m)$ if $S(x) = x$.

Opposite fixed point: An element $x \in \mathrm{GF}(2^n)$ is an opposite fixed point of S-box $S : \mathrm{GF}(2^n) \to \mathrm{GF}(2^m)$ if $S(x) = \bar{x}$.

According to [9], the S-box must not have fixed points and opposite fixed points.

Affine Equivalent: S-boxes are usually classified up to affine equivalence [1, 14, 13] since many of the immortal cryptographic properties of S-boxes, such as nonlinearity, differential uniformity, etc., are invariant under affine transformation.

Let S_1 and S_2 be $n \times n$ bit S-box. S_1 and S_2 are called affine equivalent if there exist two invertible $n \times n$ matrixes $A, B \in \mathrm{GF}(2)$, and constraints $a, b \in \mathrm{GF}(2^n)$ such that:

$$S_2 = B(S_1(A \cdot x \oplus a)) \oplus b \tag{2.2}$$

3. Algorithm for Improving Algebraic Degree of S-box Coordinate Boolean Functions

In this section, the new algorithm for improving algebraic degree of the S-box coordinate Boolean functions is presented. The proposed algorithm is based on affine transformation as presented by (2.2) and is described as follows:

Input: Original S-box (S_1) with the coordinate Boolean functions that their algebraic degrees are uneven.

Output: Affine equivalent S-box (S_1), such that the algebraic degree of coordinate Boolean functions described this S-box are uniform and higher.

Step 1: Represent n-Boolean functions of the S_1 in the form of n-ANF polynomials $\{p_1, p_2, \ldots, p_\}$ correspondingly and determine polynomial, which have the greatest degree. (We will denote the largest degree is MAXDEG.)

Step 2: Create a set Z consisting of all linear combinations and their inverse from n-original polynomials $\{p_1, p_2, \ldots, p_n\}$ of the S_1.

Thus, the set Z will include $2(2^n - 1)$ elements. They will in turn be assigned an index i ($i = 1 \ldots 2^n - 1$), respectively. That mean:

$$Z = \{z_i | z_i = i_1 p_1 \oplus i_2 p_2 \oplus \ldots \oplus i_n p_n\} \cup \{z_{2^n - 1 + i} = (1, 0, \ldots, 0) \oplus z_i\}$$

where $i_1, i_2, \ldots, i_n$ are bits of binary description of i.

Then sort Z by $\deg(z_j)$ where $j = 1,\ldots,2^(n+1)-1$.

Step 3: Choose n linearly independent polynomials and their algebraic degree in MAXDEG from the set Z.

3.1: Initial n zero vector: $v_k = 0$, for all $k = 1\ldots n$.

3.2: Initial $k = 1$ (current index of selected vector), $t = 1$.

3.3: Determine vector $z_i \in Z$, for all $i = 1\ldots 2^(n+1)-1$, such that $\deg(z_i) = \text{MAXDEG}$.

Assign $z_i \cdot a_0 = t$.

3.4: Assign $v_k = z_i$.

3.5: If $i < 2^n$ then compute:

k-th row of the matrix $B : y_k = i$

and

k-th bit of binary vector $b : b_k = 0$.

Else

k-th row of the matrix $B : y_k = i - 2^n$

and

k-th bit of the binary vector $b : b_k = 1$.

3.6: Delete vector $v_k = z_i$ and its inverse, also all linear combination from k selected vectors and their inverses.

- Linear combinations from k selected vectors and their inversions are computed as follow:

$$\text{for all } m = (m_1 m_2 \ldots m_k), \quad 2^{k-1} \le m < 2^k$$
$$x = m_1 v_1 \oplus m_2 v_2 \oplus \ldots \oplus m_k v_k$$
$$w = x \oplus \underbrace{(1,0,0,\ldots,0)}_{2^n-1}$$

  where w is the inverse of the x.
- Delete x, w from $Z : Z = Z \setminus \{x, w\}$

3.7: $k = k + 1$. If $k < n$ then $t = 0$, else $t = 1$.

3.8: If $k < n$ then go to step 3.3, else go to Step 4.

Step 4: Convert n selected polynomials in Step 3 to form of the Boolean functions in order to receive the new S-box S_2.

In this algorithm, after step 3, new S-box S_2 will be created. In this process, n vectors $(z_{y_1}, z_{y_2}, \ldots, z_{y_n})$ are determined. In other words, index set $(y_1, y_2, \ldots, y_n)$ of the component Boolean functions is selected from the set of the all combination linear and their inversions,

where $y_k \in \{1,2,\ldots,2(2^n-1)\}$, for all $k = 1,2,\ldots,n$, i.e. nonsingular matrix B and vector b have been determined, such that:

$$S_2 = BS_1 + b \tag{3.1}$$

and compared to S_1, the coordinate Boolean functions of S_2 have equal algebraic degrees and are equal to MAXDEG, where

- S_1 is the original S-box,
- B is the nonsingular matrix, which is defined as follow:

$$B = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{bmatrix} \tag{3.2}$$

  where $(y_{i1}, y_{i2}, \ldots, y_{in})$ is the binary representation of index y_i, $i = 1,2,\ldots,n$, and the value y_i is determined in step 3.5,
- $b = (b_1, b_2, \ldots, b_n)$, where $b_i \in \mathrm{GF}(2)$, $i = 1,2,\ldots,n$. The values b_i are determined in step 3.5 of this algorithm.

The (3.1) show that the new S-box (S_2) and original S-box (S_1) are affine equivalent by output.

4. Experimental Results

Applying the developed algorithm for S-box 4×4 in PRESENT (see Table 1), received results as follows:

(1) A system of S-block Boolean functions is written in the form of a truth table. Then, by multiplying a $2^n \times 2^n$ Hadamard matrix for every Boolean function compute coefficients of the ANFs recorded in the form of vectors $\{p_1, p_2, \ldots, p_n\}$ as follows and MAXDEG = 3:

p_1 : 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0

$x_0 \oplus x_2 \oplus x_1x_2 \oplus x_3$

p_2 : 0 0 1 0 0 0 0 1 1 0 1 1 1 1 0 0

$x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$

p_3 : 1 0 0 1 1 0 0 0 1 1 1 1 0 1 0 0

$1 \oplus x_0x_1 \oplus x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$

p_4 : 1 1 1 0 0 0 1 1 1 0 0 1 0 1 0 0

$1 \oplus x_0 \oplus x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$

(2) The set received Z is showed in Table 3.

Table 3. The set Z computed and sorted by proposed algorithm

i	Coefficient vectors and corresponding polynomials of ANF (z_i)		deg
30	z_{30}	1 0 0 1 0 0 0 0 0 1 0 1 1 1 0 0 $1 \oplus x_0x_1 \oplus x_0x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3

(*Contd.*)

i	Coefficient vectors and corresponding polynomials of ANF (z_i)		deg
15	z_{15}	0 0 0 1 0 0 0 0 0 1 0 1 1 1 0 0 $x_0x_1 \oplus x_0x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3
29	z_{29}	1 1 0 1 1 0 1 0 1 1 0 1 1 1 0 0 $1 \oplus x_0 \oplus x_0x_1 \oplus x_2 \oplus x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3
14	z_{14}	0 1 0 1 1 0 1 0 1 1 0 1 1 1 0 0 $x_0 \oplus x_0x_1 \oplus x_2 \oplus x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3
28	z_{28}	1 0 1 1 0 0 0 1 1 1 1 0 0 0 0 0 $1 \oplus x_1 \oplus x_0x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3$	3
13	z_{13}	0 0 1 1 0 0 0 1 1 1 1 0 0 0 0 0 $x_1 \oplus x_0x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3$	3
27	z_{27}	1 1 1 1 1 0 1 1 0 1 1 0 0 0 0 0 $1 \oplus x_0 \oplus x_1 \oplus x_0x_1 \oplus x_2 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_0x_3 \oplus x_1x_3$	3
12	z_{12}	0 1 1 1 1 0 1 1 0 1 1 0 0 0 0 0 $x_0 \oplus x_1 \oplus x_0x_1 \oplus x_2 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_0x_3 \oplus x_1x_3$	3
24	z_{24}	0 0 1 0 1 0 0 1 0 0 0 1 0 1 0 0 $x_1 \oplus x_2 \oplus x_0x_1x_2 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
9	z_9	1 0 1 0 1 0 0 1 0 0 0 1 0 1 0 0 $1 \oplus x_1 \oplus x_2 \oplus x_0x_1x_2 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
23	z_{23}	0 1 1 0 0 0 1 1 1 0 0 1 0 1 0 0 $x_0 \oplus x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
8	z_8	1 1 1 0 0 0 1 1 1 0 0 1 0 1 0 0 $1 \oplus x_0 \oplus x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
22	z_{22}	0 1 1 1 0 0 1 1 1 1 0 0 1 0 0 0 $x_0 \oplus x_1 \oplus x_0x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_2x_3$	3
7	z_7	1 1 1 1 0 0 1 1 1 1 0 0 1 0 0 0 $1 \oplus x_0 \oplus x_1 \oplus x_0x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_2x_3$	3
21	z_{21}	0 0 1 1 1 0 0 1 0 1 0 0 1 0 0 0 $x_1 \oplus x_0x_1 \oplus x_2 \oplus x_0x_1x_3 \oplus x_0x_3 \oplus x_2x_3$	3
6	z_6	1 0 1 1 1 0 0 1 0 1 0 0 1 0 0 0 $1 \oplus x_1 \oplus x_0x_1 \oplus x_2 \oplus x_0x_1x_3 \oplus x_0x_3 \oplus x_2x_3$	3
20	z_{20}	0 1 0 1 0 0 1 0 0 1 1 1 0 1 0 0 $x_0 \oplus x_0x_1 \oplus x_1x_2 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
5	z_5	1 1 0 1 0 0 1 0 0 1 1 1 0 1 0 0 $1 \oplus x_0 \oplus x_0x_1 \oplus x_1x_2 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
19	z_{19}	0 0 0 1 1 0 0 0 1 1 1 1 0 1 0 0 $x_0x_1 \oplus x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
4	z_4	1 0 0 1 1 0 0 0 1 1 1 1 0 1 0 0 $1 \oplus x_0x_1 \oplus x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
18	z_{18}	1 1 1 0 1 0 1 1 0 0 1 1 1 1 0 0 $1 \oplus x_0 \oplus x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
3	z_3	0 1 1 0 1 0 1 1 0 0 1 1 1 1 0 0 $x_0 \oplus x_1 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
17	z_{17}	1 0 1 0 0 0 0 1 1 0 1 1 1 1 0 0 $1 \oplus x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3

(*Contd.*)

i	Coefficient vectors and corresponding polynomials of ANF (z_i)		deg
2	z_2	0 0 1 0 0 0 0 1 1 0 1 1 1 1 0 0 $x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3
26	z_{26}	0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0 $x_2 \oplus x_3 \oplus x_1x_3 \oplus x_2x_3$	2
11	z_{11}	1 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0 $1 \oplus x_2 \oplus x_3 \oplus x_1x_3 \oplus x_2x_3$	2
25	z_{25}	0 1 0 0 0 0 1 0 0 0 1 0 1 0 0 0 $x_0 \oplus x_1x_2 \oplus x_1x_3 \oplus x_2x_3$	2
10	z_{10}	1 1 0 0 0 0 1 0 0 0 1 0 1 0 0 0 $1 \oplus x_0 \oplus x_1x_2 \oplus x_1x_3 \oplus x_2x_3$	2
16	z_{16}	1 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0 $1 \oplus x_0 \oplus x_2 \oplus x_1x_2 \oplus x_3$	2
1	z_1	0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0 $x_0 \oplus x_2 \oplus x_1x_2 \oplus x_3$	2

(3) Determine n-polynomials, their algebraic degree is 3. T he corresponding indexes of these polynomials are: 30, 14, 13, 9. It means that we determined the matrix B and the vector bof the affine equivalent transformation as follows:

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

and vector b

$$b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The received S-box as in Table 4:

Table 4. New affine equivalence S-box (S_2)

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S_2(x)$	8	C	B	1	D	9	A	3	E	4	F	6	7	0	2	5

The Boolean functions in ANF of this S-box are presented in Table 5.
The coordinate Boolean functions of S_2 in ANF Algebraic degree.

Table 5. Analysis of coordinate Boolean functions of the affine equivalence S-box (S_2)

$f_1(x) = 1 \oplus x_1 \oplus x_2 \oplus x_0x_1x_2 \oplus x_0x_1x_3 \oplus x_0x_2x_3$	3
$f_2(x) = x_1 \oplus x_0x_1 \oplus x_0x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_1x_3$	3
$f_3(x) = x_0 \oplus x_0x_1 \oplus x_2 \oplus x_1x_2 \oplus x_3 \oplus x_0x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3
$f_4(x) = 1 \oplus x_0x_1 \oplus x_0x_3 \oplus x_0x_1x_3 \oplus x_2x_3 \oplus x_0x_2x_3$	3

Obviously, the algebraic degrees of all Boolean functions of the received S-box are 3. The comparison of the some cryptography properties of the S-boxes S_1 and S_2 is presented in Table 6.

Table 6. The comparison of the cryptography properties of S-boxes

Property	S_1	S_2
Nonlinearity	4	4
Differential uniformity	4	4
Linear Approximation	4	4
Number of fixed points	0	0
Number of opposite fixed points	1	0
Number of coordinate Boolean functions with algebraic degree 2	1	0
Number of coordinate Boolean functions with algebraic degree 3	3	4

The comparison of the some cryptography properties of the coordinate Boolean functions of S-boxes S_1 and S_2 is presented in Table 7.

Table 7. The comparison of the some cryptography properties of the coordinate Boolean functions

The coordinate Boolean functions	S_1				S_2			
Property	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
Nonlinearity	4	4	4	4	4	4	4	4
Algebraic degree	2	3	3	3	3	3	3	3
Number of nonzero spectral Walsh coefficients	4	10	10	10	10	10	10	10
Number of nonzero spectral AC coefficients	4	7	7	7	7	7	7	7

5. Conclusion

In this paper we presented an algorithm, it allows to improve the algebraic degree of the S-box coordinate Boolean functions from applying affine equivalence transformation. By using the proposed algorithm, can create a new affine equivalent S-box with higher and more evenly algebraic degree of the coordinate Boolean functions, which may improve the ability to resist many attacks, for instance higher-order differential attacks, algebraic attacks or cube attacks. In addition to improved algebraic degree, other criteria of the S-box (for example,number of fixed points, number of opposite fixed points, number of nonzero spectral Walsh coefficients and number of nonzero spectral AC coefficients) can be added to this algorithm to further enhance the safety of the block cipher.

Acknowledgment

This article is partly supported by VAST01.06/15-16 Vietnam Academy of Science and Technology project.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References

[1] A. Biryukov, C. De Canniere, A. Braeken and B. Preneel, A toolbox for cryptanalysis: Linear and affine equivalence algorithms, InInternational Conference on the Theory and Applications of Cryptographic Techniques, Springer, Berlin — Heidelberg, pp. 33-50 (May, 2003).

[2] A. Bogdanov, L.R. Knudsen, G. Leander, C. Paar, A. Poschmann, M.J. Robshaw and C. Vikkelsoe, PRESENT: An ultra-lightweight block cipher, in *International Workshop on Cryptographic Hardware and Embedded Systems* (pp. 450-466), Springer, Berlin — Heidelberg (2007, September).

[3] N.P. Borisenko and H.D. Tho, Algorithms for minimization of the number of logic elements in a circuit implementing S-box Boolean functions, in *Proceedings of 3rd Workshop on Current Trends in Cryptology* (CTCrypt'2014), Moscow, Russia (2014).

[4] C. Boura and A. Canteaut, On the influence of the algebraic degree of on the algebraic degree of. *IEEE Transactions on Information Theory* **59**(1) (2013), 691 – 702.

[5] C. Boura, A. Canteaut and C. De Canniere, Higher-order differential properties of Keccak and Luffa, in *International Workshop on Fast Software Encryption* (pp. 252-269), Springer, Berlin — Heidelberg (February, 2011).

[6] C. Carlet, Boolean functions for cryptography and error correcting codes, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering* **2** (2010), 257.

[7] Y. Crama and P.L. Hammer, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering* (Vol. 2), Cambridge University Press (2010).

[8] T.W. Cusick and P. Stanica„ *Cryptographic Boolean Functions and Applications*, Academic Press (2009).

[9] J. Daemen and V. Rijmen, *The Design of Rijndael: AES — the Advanced Encryption Standard*, Springer Science & Business Media (2013).

[10] L.T. Dung and H.D. Tho, An algorithm for improving algebraic degree of S-box based on affine equivalence transformation, *International Journal of Knowledge and Systems Science* **8**(1) (2017), 53 – 64.

[11] A. Grocholewska-Czuryło, Cryptographic properties of modified AES-like S-boxes, *Annales UMCS, Informatica* **11**(2) (January, 2011), 37 – 48.

[12] G. Ivanov, N. Nikolov and S. Nikova, Reversed genetic algorithms for generation of bijective S-boxes with good cryptographic properties, *Cryptography and Communications* **8**(2) (2016), 247 – 276.

[13] G. Leander and A. Poschmann, On the classification of 4 bit S-boxes, in: *International Workshop on the Arithmetic of Finite Fields*, pp. 159-176, Springer, Berlin — Heidelberg (June, 2007).

[14] J. McLaughlin and J.A. Clark, Using evolutionary computation to create vectorial Boolean functions with low differential uniformity and high nonlinearity, arXiv preprint arXiv:1301.6972 (2013).

[15] National Institute for Science and Technology (NIST), *Advanced Encryption Standard* (AES), URL: `http://www.csrc.nist.gov/publications/fips/fips197/fips-197.pdf` (November, 2001).

[16] National Institute for Science andTechnology (NIST), *Data Encryption Standard* (DES), URL: `http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf` (October, 1999).

[17] B. Preneel and A. Braeken, *Cryptographic Properties of Boolean Functions and S-Boxes* (2006).

[18] A.G. Rostovtsev and E.B. Mahovenko, *Theoretical cryptography*, St. Petersburg Press (in Russian) (2005).

[19] C.E. Shannon, Communication theory of secrecy systems, *Bell System Technical Journal* **28**(4) (1949), 656 – 715.

[20] Y. Tan, G. Gong and B. Zhu, Enhanced criteria on differential uniformity and nonlinearity of cryptographically significant functions, *Cryptography and Communications* **8**(2) (2016), 291 – 311.