Finding an effective metric used for bijective S-Box Generation by Genetic Algorithms

Tsonka Baicheva, Dusan Bikov, Yuri Borissov, Limonka Lazarova, Aleksandra Stojanova, Liliya Stoykova, Stela Zhelezova

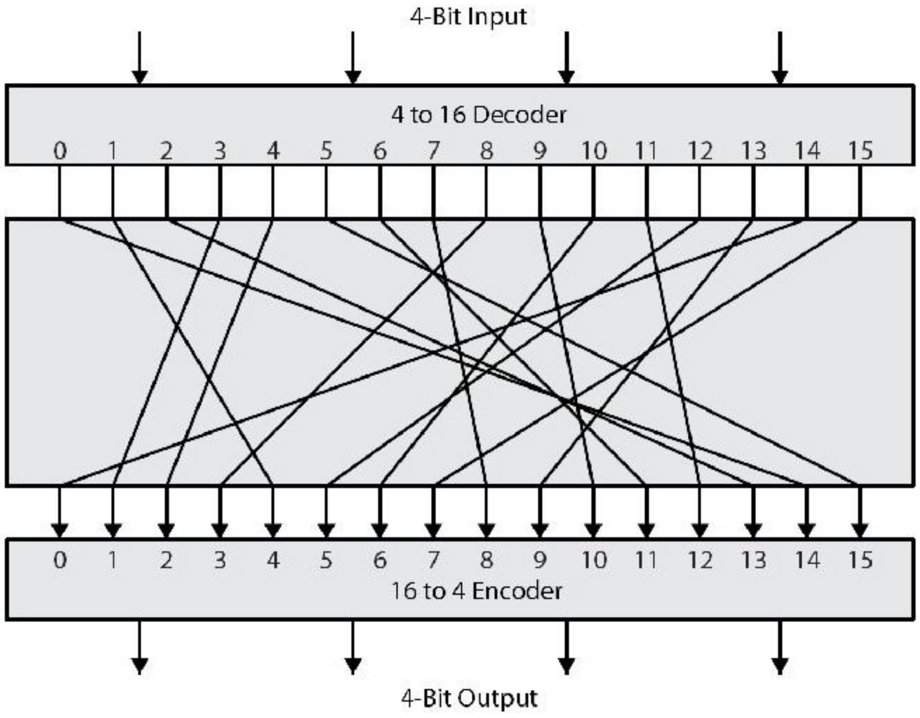
Problem set by SANS: Nikolay Nikolov, Georgi Ivanov

27.09.2014

WHAT IS S-BOX?

In cryptography, an S-box (substitution-box) is a basic component of symmetric key algorithms which performs nonlinear substitution.

S-boxes transform n-binary input into m-binary output.



What is Boolean function?

- Let B={0,1} and $B^n = \{0, 1\}^n$. Every function $f : B^n \to B$ is called Boolean function of n variables.

$$\boldsymbol{B}_n = \{ f \mid f : B^n \to B \},\ |\boldsymbol{B}_n| = 2^{2^n}$$

- Let $f_1, f_2, \ldots, f_m \in \boldsymbol{B}_n$. Mapping $F : B^n \to B^m$ defined by the rule $F(x) = (f_1(x), f_2(x), \ldots, f_m(x))$, is called vectorial Boolean function and $f_1, f_2, \ldots, f_m$ are its coordinate functions.

What is S-box?

Let S be the substitution table of an **n-binary input**

into **m-binary output** mapping, that is, if $B = \{0,1\}$.

$$S : B^n \rightarrow B^m$$

$$\mathbf{x = (x_1, x_2, \ldots, x_n) \rightarrow y = (y_1, y_2 \ldots, y_m) = S(x)}$$

S can be considered as a vectorial Boolean function,

consisting of **m** individual **n**-variable Boolean functions

$f_1, f_2, \ldots, f_m$, referred to as **coordinate** Boolean functions

Bijective S-Box

- An (n x n) S-Box **S** is called bijective, if **S** is an invertible mapping over B^n .

- Bijective S-Boxes represent **permutations** of their 2^n inputs.

- Walsh-Hadamard Transform (WHT) spectrum of f(x) is the set of all 2^n spectral coefficients for the elements in B^n

- WHT Spectrum Matrix is the matrix of WHT spectrum of all coordinate Boolean functions.

S-Box criteria

The main criteria for cryptographically strong (n x n)

S-Box

- High nonlinearity
- High algebraic degree
- Balanced structure
- Good autocorrelation properties

Bent S-Box

An (n x m) S-Box **S** is referred as a **Bent** S-Box, if WHT Spectrum Matrix is entirely flat.

Bent S-Box has the highest possible nonlinearity.

Bent itself is not satisfying for our purposes- It is not balanced and exists only for even n $\geq$ 2m.

The Group Goal

We focus on achieving good performance according to the nonlinearity criterion – finding S-Box close to Bent S-Box.

Our suggestions

Suggestion 1: Exponential S-Box

For the genetic algorithm, generate the initial parent pool of bijective S-Boxes, $\mathbf{P_1}$, $\mathbf{P_2}$, … , $\mathbf{P_T}$, where $\mathbf{P_i}$ are **exponential S-boxes**. (S. Agievich, A. Afonenko, 2005)

Exponential S-Boxes are proven to have good cryptographic properties.

Our suggestions

Suggestion 2: Change the cost function

The cost function can be computed by using the maximum of the differences between the spectral coefficients of each coordinate function.

S-Box

$b_{0,0}$	$b_{1,0}$		$b_{2^n-1,0}$
$b_{0,1}$	$b_{1,1}$		$b_{2^n-1,1}$
. . .	. . .	. . .	
$b_{0,2^n-1}$	$b_{1,2^n-1}$		$b_{2^n-1,2^n-1}$

WHT spectrum

$w_{o,o}$	$w_{1,o}$		$w_{o,o}$
$w_{o,1}$	$w_{1,1}$		$w_{1,2^n-1}$
$w_{o,2^n-1}$	$w_{o,2^n-1}$		$w_{2^n-1,2^n-1}$

Their cost function:

$$\sqrt[P]{\sum_{j=0}^{2^n-1} \left|w_{i,j} - w_{i,j+1}\right|^P}$$

Our cost function:

Δ_i-max difference of i[th] coordinate

$$\Delta_i = max_i \begin{Bmatrix} w_{i,j} - w_{i,j+k} | j \in (0, 2^n - 1); \\ j + k \le 2^n - 1 \end{Bmatrix}$$

$$\Delta_{WHT} = max\Delta_i, i \in (1, 2^n - 1)$$

1) $\Delta_{BENT} = 0$

2) $\begin{matrix} (\Delta_1, \Delta_2, \dots, \Delta_{2^n-1})\ S_1 \\ (\Delta_1, \Delta_2, \dots, \Delta_{2^n-1})\ S_2 \end{matrix} \rightarrow count\ \Delta_i = 0$

If $\Delta_{S_1} \approx \Delta_{S_2}$ then the second condition can be used and the S-box which has more $\Delta_i = 0$ is chosen

OUR SUGGESTIONS

Suggestion 3: Change the cost function

Another cost function can be computed by using the dispersion of the WHT spectrum of each coordinate function .

Let a_0, $a_1, \ldots,$ a_{2k} be possible values of the WHT spectrum matrix and $p_{i,j}$ be the probability of appearing a_j in the i^{th} column. Then the mathematical expectation is

$$E(w_i) = \sum_{j=0}^{2k} a_j p_{i,j}$$

OUR SUGGESTIONS

- The dispersion of the i[th] column of the WHT matrix is:

$$D(w_i) = E(w_i^2) - \left(2^{\frac{n}{2}}\right)^2 = \sum_{j=0}^{2k} a_j^2 p_{i,j} - 2^n$$

The dispersion of the S-Box is:

$$D(S) = \frac{1}{2^{n-1}} \sum_{i=1}^{2^n-1} D(w_i)$$

Smaller dispersion means flatter spectrum and better S-Box.

OUR SUGGESTIONS

Suggestion 4: Examine smaller S-Boxes

Examine the behavior of the genetic algorithm on 4x4 S-boxes and compare the results with the already known optimal ones.

- This can give verification of the method and some suggestions for the cost function.

Our suggestions

Suggestion 5: New approach

- Quasigroups as a Tool for Construction of Optimal S-boxes:
  - An algorithm for construction of optimal 4x4 S-box already exists.(H.Mihajloska, D.Gligoroski, 2012)
  - Cryptographically strong 6×4-bit, 8×8-bit and other types of S-Boxes could be produced by extending the above algorithm.

Thank you for your attention!