Advances in Mathematics of Communications
Volume X, No. 0X, 20xx, X–XX
doi:10.3934/amc.xx.xx.xx

GENERALIZED NONLINEARITY OF S-BOXES

Sugata Gangopadhyay
Department of Computer Science and Engineering,
Indian Institute of Technology Roorkee, Roorkee 247667, INDIA
sugatfma@iitr.ac.in

Goutam Paul[1]
Cryptology and Security Research Unit,
R. C. Bose Centre for Cryptology and Security,
Indian Statistical Institute, Kolkata 700108, INDIA
goutam.paul@isical.ac.in

Nishant Sinha
Department of Computer Science and Engineering,
Indian Institute of Technology Roorkee, Roorkee 247667, INDIA
nishantsinha.iitr@gmail.com

Pantelimon Stănică
Department of Applied Mathematics, Naval Postgraduate School,
Monterey, CA 93943–5216, USA
pstanica@nps.edu

(Communicated by the associate editor name)

Abstract. While analyzing S-boxes, or vectorial Boolean functions, it is of interest to approximate its component functions by affine functions. In the usual attack models, it is assumed that all input vectors to an S-box are equiprobable. The nonlinearity of an S-box is defined, subject to this assumption. In this paper, we explore the possibility of linear cryptanalysis of an S-box by introducing biased inputs and thus propose a generalized notion of nonlinearity along with a generalization of the Walsh-Hadamard spectrum of an S-box.

1. Introduction

Let $\mathbb{F}_2$ be the finite field with two elements and $\mathbb{Z}$ be the ring of integers. For any $n \in \mathbb{Z}^+$, the set of positive integers, let $[n] = \{1, \ldots, n\}$. The Cartesian product of n copies of $\mathbb{F}_2$ is $\mathbb{F}_2^n = \{\mathbf{x} = (x_1, \ldots, x_n) : x_i \in \mathbb{F}_2, i \in [n]\}$ which is an n-dimensional vector space over $\mathbb{F}_2$. For any $m, n \in \mathbb{Z}^+$, a function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is said to be an (n, m) *vectorial Boolean function* (in short, an (n, m)-function) or an $n \times m$ *S-box*. An $(n, 1)$-function is said to be a Boolean function in n variables. The S-boxes are important components in block cipher designs, since usually they are the only sources of nonlinearity. DES S-boxes have been studied for more than three decades and their analysis is still relevant today. Similarly, the AES S-box is studied extensively.

2010 *Mathematics Subject Classification:* Primary: 06E30, 11T71; Secondary: 94A60.
Key words and phrases: Nonlinearity, S-box, Vectorial Boolean function, Walsh-Hadamard transform.
Nishant Sinha thanks IIT Roorkee for supporting his research.
[1] *Corresponding Author*

©20xx AIMS

Matsui [5] introduced the linear cryptanalysis of block ciphers which involves linear approximation of the S-boxes employed in their designs. The component functions of the S-boxes are approximated by linear Boolean functions by assuming all the input vectors to be equiprobable. In this paper, we generalize the notion of linear approximation of S-boxes by introducing a framework where some of the input variables are biased although all the variables are independent. We sketch the possibility of a chosen-plaintext attack based on these considerations.

Boolean functions with biased inputs, which we refer to as μ_p-Boolean functions, is a common generalization of Boolean functions which stems from the theory of random graphs developed by [1]. The graph properties in a random graph expressed as such Boolean functions are used by Friedgut and Kalai [2]. For a detailed discussion on the Fourier analysis of μ_p-Boolean functions we refer to [6, Chapter 8]. Biased analysis is recently considered for cryptanalysis of the stream ciphers E_0 and Shannon cipher by Lu and Desmedt [4]. Generalized S-box nonlinearity has been considered by Parker [7], using nega-Hadamard spectrum of S-boxes. The recent work [3] has shown the connection between Boolean functions with biased inputs and nega-Hadamard spectra by resorting to quantum implementations of Boolean functions.

Our current work establishes a new design criteria for cryptographically secure S-boxes. We also believe that this is an important step towards developing cryptanalytic techniques based on Parker's theory [7].

2. Linear approximations of an S-box

An S-box F can also be thought of as a sequence of Boolean functions written as $F = (f_1, \ldots, f_m)$ where each $f_i : \mathbb{F}_2^n \to \mathbb{F}_2$ is a Boolean function. These are said to be *coordinate functions* of F. $\mathbb{F}_2$-linear combinations of coordinate functions are said to be component functions. For any $\mathbf{v} \in \mathbb{F}_2^m$, $\mathbf{v}\cdot F$ is a *component function* of F. The inner product is defined by $\mathbf{x}\cdot\mathbf{y} = \bigoplus_{i\in[n]} x_i y_i$. The linear function corresponding to $\mathbf{u} \in \mathbb{F}_2^n$ is $\varphi_{\mathbf{u}}(\mathbf{x}) = \mathbf{u}\cdot\mathbf{x}$, for all $\mathbf{x} \in \mathbb{F}_2^n$. The *intersection* of two vectors $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{F}_2^n$ is defined by $\mathbf{x} * \mathbf{y} = (x_1y_1, \ldots, x_ny_n)$. The (Hamming) distance between two Boolean functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ is $d(f,g) = |\{\mathbf{x} \in \mathbb{F}_2^n : f(\mathbf{x}) \neq g(\mathbf{x})\}|$. Thus, the Hamming distance between the component function $\mathbf{v}\cdot F$ and the linear function $\varphi_{\mathbf{u}}$ is

$$\begin{aligned} d(\mathbf{v}\cdot F, \varphi_{\mathbf{u}}) &= |\{\mathbf{x} \in \mathbb{F}_2^n : \mathbf{v}\cdot F(\mathbf{x}) \neq \varphi_{\mathbf{u}}(\mathbf{x})\}| \\ &= 2^{n-1} - \frac{1}{2}\sum_{\mathbf{x}\in\mathbb{F}_2^n} (-1)^{\mathbf{v}\cdot F(\mathbf{x})+\varphi_{\mathbf{u}}(\mathbf{x})} \\ &= 2^{n-1} - \frac{1}{2} W_F(\mathbf{u}, \mathbf{v}), \end{aligned} \tag{1}$$

where $W_F(\mathbf{u},\mathbf{v}) = \sum_{\mathbf{x}\in\mathbb{F}_2^n} (-1)^{\mathbf{v}\cdot F(\mathbf{x})+\varphi_{\mathbf{u}}(\mathbf{x})}$. The nonlinearity of $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is given by

$$nl(F) = \min_{\mathbf{u}\in\mathbb{F}_2^n, \mathbf{v}\in\mathbb{F}_2^m} d(\mathbf{v}\cdot F, \varphi_{\mathbf{u}}) = 2^{n-1} - \frac{1}{2} \max_{\mathbf{u}\in\mathbb{F}_2^n, \mathbf{v}\in\mathbb{F}_2^m} |W_F(\mathbf{u},\mathbf{v})|\,. \tag{2}$$

Suppose that

$$\max_{\mathbf{u}\in\mathbb{F}_2^n, \mathbf{v}\in\mathbb{F}_2^m} |W_F(\mathbf{u},\mathbf{v})| = |W_F(\mathbf{u}_0,\mathbf{v}_0)|\,.$$

If the value of $|W_F(\mathbf{u}_0,\mathbf{v}_0)|$ is high then the component function $\mathbf{v}_0\cdot F(\mathbf{x})$ can be efficiently approximated by $\mathbf{u}_0\cdot\mathbf{x}$ or its complement.

Suppose that a plaintext block $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$ XOR-ed to the key $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{F}_2^n$ after being acted upon by F produces the ciphertext $\mathbf{y} = (y_1, \ldots, y_m) = F(\mathbf{x} \oplus \mathbf{k}) \in \mathbb{F}_2^m$. Let $\mathbf{X}, \mathbf{Y}, \mathbf{K}$ be the random variables corresponding to plaintext, ciphertext and key, respectively. Then

$$\begin{aligned} \Pr[\mathbf{u}_0 \cdot \mathbf{K} = \mathbf{v}_0 \cdot \mathbf{Y} \oplus \mathbf{u}_0 \cdot \mathbf{X}] &= \Pr[\mathbf{u}_0 \cdot (\mathbf{X} \oplus \mathbf{K}) = \mathbf{v}_0 \cdot \mathbf{Y}] \\ &= 1 - \frac{2^{n-1} - \frac{1}{2} W_F(\mathbf{u}_0, \mathbf{v}_0)}{2^n} \\ &= \frac{2^{n-1} + \frac{1}{2} W_F(\mathbf{u}_0, \mathbf{v}_0)}{2^n} \\ &= \frac{1}{2} + \frac{1}{2^{n+1}} W_F(\mathbf{u}_0, \mathbf{v}_0). \end{aligned} \tag{3}$$

Thus, if $W_F(\mathbf{u}_0, \mathbf{v}_0) \geq 0$ and $2^{-(n+1)} W_F(\mathbf{u}_0, \mathbf{v}_0)$ is close to $\frac{1}{2}$ then the equation $\mathbf{u}_0 \cdot \mathbf{K} = \mathbf{v}_0 \cdot \mathbf{Y} \oplus \mathbf{u}_0 \cdot \mathbf{X}$ is true with a probability close to 1. Similarly, if $W_F(\mathbf{u}_0, \mathbf{v}_0) \leq 0$ and $2^{-(n+1)} W_F(\mathbf{u}_0, \mathbf{v}_0)$ is close to $-\frac{1}{2}$ then the equation $\mathbf{u}_0 \cdot \mathbf{K} = \mathbf{v}_0 \cdot \mathbf{Y} \oplus \mathbf{u}_0 \cdot \mathbf{X} \oplus 1$ is true with a probability close to 1. Thus if we have a large sample of the plaintext-ciphertext pairs we will be able to establish linear relationships between the key bits. This leads to linear cryptanalysis of block ciphers which was introduced by [5] for cryptanalysis of DES.

3. Linear approximation of an S-box with respect to partially biased inputs

Let $\mathcal{S} \subseteq [n]$. $\mathbf{X} = (X_1, \ldots, X_n)$, and $\mathbf{K} = (K_1, \ldots, K_n)$ are n-tuples and $\mathbf{Y} = (Y_1, \ldots, Y_m)$ is an m-tuple of random variables corresponding to plaintexts, chipertexts and keys, respectively, such that

$$\mathbf{Y} = F(\mathbf{X} \oplus \mathbf{K}) \tag{4}$$

where $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$.

3.1. Linear approximations when the inputs are partially biased.

To simplify the notation assume that the input to $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is $\mathbf{X} = (X_1, \ldots, X_n)$ having the following distribution.

$$\Pr[X_i = 1] = \begin{cases} p, & \text{if } i \in \mathcal{S} \\ \frac{1}{2}, & \text{if } i \in [n] \setminus \mathcal{S}, \end{cases} \qquad \Pr[X_i = 0] = \begin{cases} 1 - p, & \text{if } i \in \mathcal{S} \\ \frac{1}{2}, & \text{if } i \in [n] \setminus \mathcal{S}. \end{cases} \tag{5}$$

We say that a component function $\mathbf{v} \cdot F$ can be approximated by $\varphi_{\mathbf{u}}$ if $\Pr[\mathbf{v} \cdot F(\mathbf{X}) = \varphi_{\mathbf{u}}(\mathbf{X})]$ is high. Let $\mathbf{e}_i \in \mathbb{F}_2^n$ be the vector whose ith component is 1 and remaining components are 0's, for all $i \in [n]$. Define $\mathbf{e}_{\mathcal{S}} = \bigoplus_{i \in \mathcal{S}} \mathbf{e}_i$. We introduce a notion of

distance similar to that in [3] as follows:

$$
\begin{aligned}
& d_{\mathcal{S}}^{(p)}(\mathbf{v}\cdot F, \varphi_{\mathbf{u}}) \\
&= 2^n \Pr[\mathbf{v}\cdot F(\mathbf{X}) \neq \varphi_{\mathbf{u}}(\mathbf{X})] \\
&= 2^n \sum_{\mathbf{v}\cdot F(\mathbf{x})\neq \mathbf{u}\cdot\mathbf{x}} 2^{-(n-|\mathcal{S}|)} p^{\mathrm{wt}(\mathbf{e}_{\mathcal{S}}\cdot\mathbf{x})}(1-p)^{|\mathcal{S}|-\mathrm{wt}(\mathbf{e}_{\mathcal{S}}\cdot\mathbf{x})} \\
&= 2^{|\mathcal{S}|-2}(1-p)^{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathbb{F}_2^n} \left(\frac{p}{1-p}\right)^{\mathrm{wt}(\mathbf{e}_{\mathcal{S}}\cdot\mathbf{x})} ((-1)^{\mathbf{v}\cdot F(\mathbf{x})} - (-1)^{\mathbf{u}\cdot\mathbf{x}})^2 \\
&= 2^{|\mathcal{S}|-2}(1-p)^{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathbb{F}_2^n} \left(\frac{p}{1-p}\right)^{\mathrm{wt}(\mathbf{e}_{\mathcal{S}}\cdot\mathbf{x})} (2 - 2(-1)^{\mathbf{v}\cdot F(\mathbf{x})\oplus\mathbf{u}\cdot\mathbf{x}}) \qquad (6) \\
&= 2^{|\mathcal{S}|-1}(1-p)^{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathbb{F}_2^n} \left(\frac{p}{1-p}\right)^{\mathrm{wt}(\mathbf{e}_{\mathcal{S}}\cdot\mathbf{x})} - \frac{2^{|\mathcal{S}|}(1-p)^{|\mathcal{S}|}}{2} W_{F,\mathcal{S}}^{(p)}(\mathbf{u},\mathbf{v}) \\
&= 2^{|\mathcal{S}|-1}(1-p)^{|\mathcal{S}|} 2^{n-|\mathcal{S}|}(1-p)^{-|\mathcal{S}|} - \frac{2^{|\mathcal{S}|}(1-p)^{|\mathcal{S}|}}{2} W_{F,\mathcal{S}}^{(p)}(\mathbf{u},\mathbf{v}) \\
&= 2^{n-1} - \frac{2^{|\mathcal{S}|}(1-p)^{|\mathcal{S}|}}{2} W_{F,\mathcal{S}}^{(p)}(\mathbf{u},\mathbf{v})
\end{aligned}
$$

where

$$
W_{F,\mathcal{S}}^{(p)}(\mathbf{u},\mathbf{v}) = \sum_{\mathbf{x}\in\mathbb{F}_2^n} (-1)^{\mathbf{v}\cdot F(\mathbf{x})\oplus\mathbf{u}\cdot\mathbf{x}} \left(\frac{p}{1-p}\right)^{\mathrm{wt}(\mathbf{e}_{\mathcal{S}}\cdot\mathbf{x})} \qquad (7)
$$

is a generalization of the Walsh–Hadamard transform for vectorial Boolean functions (S-boxes). As a side problem, we propose to the community that the transform defined in (7) be investigated for different S-boxes.

3.2. A CHOSEN-PLAINTEXT ATTACK MODEL. Let the coordinates of $\mathbf{K}$ be *i.i.d.* (independent and identically distributed), i.e.,

$$
\Pr[K_i = 1] = \Pr[K_i = 0] = \frac{1}{2}, \text{ for all } i \in [n]. \qquad (8)
$$

If X_i and K_i had been independent, then assuming $\Pr[X_i = 0] = q_i$, we would have

$$
\begin{aligned}
& \Pr[X_i \oplus K_i = 0] \\
&= \Pr[X_i = 0, K_i = 0] + \Pr[X_i = 1, K_i = 1] \\
&= \Pr[X_i = 0]\cdot\Pr[K_i = 0] + \Pr[X_i = 1]\cdot\Pr[K_i = 1] \\
&= q_i \cdot \frac{1}{2} + (1-q_i)\cdot\frac{1}{2} \\
&= \frac{1}{2}.
\end{aligned}
$$

In other words, the distribution of $X_i \oplus K_i$ would have been unbiased, irrespective of the bias in X_i. But since we make X_i dependent on our guess of K_i, the above result do not hold and we can bias the distribution of $X_i \oplus K_i$. If we guess the

key-bits k_i for all $i \in \mathcal{S}$, then for all $i \in [n]$ we can simulate $X_i \oplus K_i$ such that

$$\begin{aligned}\Pr[X_i \oplus K_i = 0] &= \begin{cases} 1-p, & \text{if } i \in \mathcal{S} \\ \frac{1}{2}, & \text{if } i \in [n] \setminus \mathcal{S}, \end{cases} \\ \Pr[X_i \oplus K_i = 1] &= \begin{cases} p, & \text{if } i \in \mathcal{S} \\ \frac{1}{2}, & \text{if } i \in [n] \setminus \mathcal{S}, \end{cases}\end{aligned} \tag{9}$$

for any $0 \leq p \leq 1$, by choosing X_i's appropriately, and ensuring that $X_i \oplus K_i$ are independent random variables.

From the discussion above we observe that employing a chosen-plaintext model it is possible to ensure that the input to an S-box (equivalently, the vectorial Boolean function) F is partially-biased. With respect to such a partially-biased input if a component function of F can be approximated by a linear function $\varphi_{\mathbf{u}}$ or its complement then one may be able to apply linear cryptanalysis techniques on F.

Suppose that $\mathbf{v}_0 \cdot F$ is close to a linear function $\varphi_{\mathbf{u}}$ when the inputs of the variables with indexes in $\mathcal{S}_0$ are biased. Then we can choose plaintexts $\mathbf{x}$ with respect to each guess of the key segment $\mathbf{e}_{\mathcal{S}_0} * \mathbf{k}$ such that the following equation is satisfied with high probability

$$\begin{aligned} & \mathbf{u}_0 \cdot (\mathbf{x} \oplus \mathbf{k}) = \mathbf{v}_0 \cdot \mathbf{y} \\ \textit{i.e.}, \ & \mathbf{u}_0 \cdot \mathbf{k} = \mathbf{u}_0 \cdot \mathbf{x} \oplus \mathbf{v}_0 \cdot \mathbf{y} \\ \textit{i.e.}, \ & \mathbf{u}_0 \cdot (\mathbf{e}_{\mathcal{S}_0} * \mathbf{k} \oplus (1 \oplus \mathbf{e}_{\mathcal{S}_0}) * \mathbf{k}) = \mathbf{u}_0 \cdot \mathbf{x} \oplus \mathbf{v}_0 \cdot \mathbf{y} \\ \textit{i.e.}, \ & \mathbf{u}_0 \cdot ((1 \oplus \mathbf{e}_{\mathcal{S}_0}) * \mathbf{k}) = \mathbf{u}_0 \cdot (\mathbf{e}_{\mathcal{S}_0} * \mathbf{k}) \oplus \mathbf{u}_0 \cdot \mathbf{x} \oplus \mathbf{v}_0 \cdot \mathbf{y} \end{aligned}$$

To be more precise

$$\begin{aligned} & \Pr[\mathbf{u}_0 \cdot ((1 \oplus \mathbf{e}_{\mathcal{S}_0}) * \mathbf{K}) = \mathbf{u}_0 \cdot (\mathbf{e}_{\mathcal{S}_0} * \mathbf{K}) \oplus \mathbf{u}_0 \cdot \mathbf{X} \oplus \mathbf{v}_0 \cdot \mathbf{Y}] \\ &= 1 - \frac{2^{n-1} - \frac{2^{|\mathcal{S}_0|}(1-p)^{|\mathcal{S}_0|}}{2} W^{(p)}_{F,\mathcal{S}_0}(\mathbf{u}_0, \mathbf{v}_0)}{2^n} \\ &= \frac{1}{2} + \frac{2^{|\mathcal{S}_0|}(1-p)^{|\mathcal{S}_0|}}{2^{n+1}} W^{(p)}_{F,\mathcal{S}_0}(\mathbf{u}_0, \mathbf{v}_0) \end{aligned} \tag{10}$$

Thus the knowledge of $\mathbf{u}_0 \cdot (\mathbf{e}_{\mathcal{S}_0} * \mathbf{k})$, $\mathbf{u}_0 \cdot \mathbf{x}$ and $\mathbf{v}_0 \cdot \mathbf{y}$ allows us to derive relationships between the unknown key-bits of $\mathbf{k}$ which form the vector $(1 \oplus \mathbf{e}_{\mathcal{S}_0}) * \mathbf{k}$. This might be translated to key recovery in time less than the exhaustive search.

3.3. MOUNTING A MORE EFFICIENT LINEAR CRYPTANALYSIS. Let $\mathcal{S} \subseteq [n]$, $\mathbf{u} \in \mathbb{F}_2^n$ and $\mathbf{v} \in \mathbb{F}_2^m$. Consider an S-box $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$. Then from Equation (3), we have

$$\Pr[\mathbf{u} \cdot \mathbf{K} = \mathbf{v} \cdot \mathbf{Y} \oplus \mathbf{u} \cdot \mathbf{X}] = \frac{1}{2} + \frac{1}{2^{n+1}} W_F(\mathbf{u}, \mathbf{v}).$$

The absolute value of the bias of the above event is

$$\epsilon(\mathbf{u}, \mathbf{v} \cdot F) = \frac{1}{2^{n+1}} \left| W_F(\mathbf{u}, \mathbf{v}) \right|.$$

If the inputs follow the probability distribution (5), then from Equation (10), we have

$$\Pr[\mathbf{u} \cdot ((1 \oplus \mathbf{e}_{\mathcal{S}}) * \mathbf{K}) = \mathbf{u} \cdot (\mathbf{e}_{\mathcal{S}} * \mathbf{K}) \oplus \mathbf{u} \cdot \mathbf{X} \oplus \mathbf{v} \cdot \mathbf{Y}] = \frac{1}{2} + \frac{2^{|\mathcal{S}|}(1-p)^{|\mathcal{S}|}}{2^{n+1}} W^{(p)}_{F,\mathcal{S}}(\mathbf{u}, \mathbf{v})$$

and the corresponding absolute value of the bias

$$\epsilon^{(p)}_{\mathcal{S}}(\mathbf{u}, \mathbf{v} \cdot F) = \frac{2^{|\mathcal{S}|}(1-p)^{|\mathcal{S}|}}{2^{n+1}} \left| W^{(p)}_{F,\mathcal{S}}(\mathbf{u}, \mathbf{v}) \right|.$$

According to [5], if for an SPN-based iterated block cipher we can pile-up the biases up to the last-but-one round, and we can form a linear equation involving a subset of the key bits with a probability $q \neq \frac{1}{2}$, then we can mount linear cryptanalysis to find the values of that subset of key bits independent of the other key bits. For a constant success probability, the number of samples (i.e., plaintext-ciphertext pairs) required is given by $\frac{1}{\left(q-\frac{1}{2}\right)^2} = \frac{1}{\delta^2}$, where $\delta = |q - \frac{1}{2}|$.

Now, for some S-box F used in the block cipher, for some $\mathcal{S} \in [n]$, if we can find a suitable p, $0 < p < 1$ and $\mathbf{v} \in \mathbb{F}_2^m$, such that

$$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\mathcal{S}}^{(p)}(\mathbf{u}, \mathbf{v}\cdot F) > \max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{v}\cdot F), \tag{11}$$

then we can pile-up the biases such that the resulting bias q_p (for biased inputs as per distribution (5)) of the subkey-dependent expression corresponding to the last-but-one round is larger than the usual bias q without biased input. Thus, $\delta_p = |q_p - \frac{1}{2}|$ would be greater than δ, thereby requiring less number of samples for the linear cryptanalysis using biased inputs. The reduced data complexity would also lead to reduced time complexity of the attack.

Note that, as long as one is able to find at least one $\mathcal{S} \in [n]$, one suitable p, $0 < p < 1$ such that Equation (11) holds, then one can mount a better attack. However, if one performs an offline exhaustive enumeration of all the biases by varying all the parameters, then one would be able to mount the optimal attack.

3.4. A PILING-UP LEMMA IN THE BIASED CONTEXT. Let $\mathcal{S} \subseteq [n]$ and X_i, $1 \leq i \leq k$, be independent random variables whose values are as in (5). We let

$$p_i = \begin{cases} p & \text{if } i \in \mathcal{S} \\ 1/2 & \text{if } i \notin \mathcal{S}, \end{cases}$$

and the *bias* of a random variable X with some probability distribution p is defined by $\epsilon_X = p - \frac{1}{2}$. Thus, for our X_i,

$$\epsilon_{X_i} = p_i - \frac{1}{2} = \begin{cases} p - \frac{1}{2} & \text{if } i \in \mathcal{S} \\ 0 & \text{if } i \notin \mathcal{S} \end{cases}$$

(we shall often write ϵ_i in lieu of ϵ_{X_i}, if it is clear from the context).

Lemma 3.1 (**Piling-up Lemma** [8, p. 81])**.** *Let $\mathcal{S} \subseteq [n]$ and X_i, $1 \leq i \leq k$, be independent random variables whose values are as in (5). Then the probability that $X_1 \oplus X_2 \oplus \cdots \oplus X_k = 0$ is*

$$\Pr(X_1 \oplus X_2 \oplus \cdots \oplus X_k = 0) = \frac{1}{2} + 2^{k-1} \prod_{i=1}^{k} \epsilon_i,$$

and therefore, the bias for $X = X_1 \oplus X_2 \oplus \cdots \oplus X_k$ is $\epsilon_X = 2^{k-1} \prod_{i=1}^{k} \epsilon_i$.

The following corollary is obvious.

Corollary 1. $\Pr(X_1 \oplus X_2 \oplus \cdots \oplus X_k = 0) = \frac{1}{2}$ *if and only if there exists $1 \leq i \leq k$ with $i \notin \mathcal{S}$.*

4. EXPERIMENTAL RESULTS WITH DES AND AES S-BOXES

In Section 3.3, we have discussed how to mount a more efficient attack with biased inputs, if we can have

$$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\mathcal{S}}^{(p)}(\mathbf{u}, \mathbf{v}\cdot F) > \max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{v}\cdot F)$$

for some S-box F used in a block cipher. In this section, we analyse some S-boxes F by computing the following ordered pair

$$(\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{v}\cdot F), \max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\mathcal{S}}^{(p)}(\mathbf{u}, \mathbf{v}\cdot F)).$$

and listing them in a table.

4.1. EXPERIMENTS WITH DES S-BOX.

Let F be the first DES S-box. The truth table of $\mathbf{e}_1 \cdot F$ is

$$(1001100001101110011001110110000101011110100100101011100101100001).$$

From the direct computation we observe that

$$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{e}_1\cdot F) \approx 0.219,$$

whereas

$$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\{2,4,6\}}^{(0.99)}(\mathbf{u}, \mathbf{e}_1\cdot F) \approx 0.494. \tag{12}$$

Suppose that $\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\{2,4,6\}}^{(0.99)}(\mathbf{u}, \mathbf{e}_1\cdot F) = \epsilon_{\{2,4,6\}}^{(0.99)}(\mathbf{u}_0, \mathbf{e}_1\cdot F)$. Therefore,

$$\Pr[\mathbf{u}_0\cdot((1\oplus \mathbf{e}_{\{2,4,6\}})*\mathbf{K}) = \mathbf{u}_0\cdot(\mathbf{e}_{\{2,4,6\}}*\mathbf{K}) \oplus \mathbf{u}_0\cdot\mathbf{X}\oplus\mathbf{v}\cdot\mathbf{Y}] \in \{0.994, 0.006\}.$$

It seems that if we assume values of $\mathbf{e}_{\{2,4,6\}}*\mathbf{K}$ and choose $\mathbf{X}$ as described in the attack model, then the either

$$\mathbf{u}_0\cdot((1\oplus \mathbf{e}_{\{2,4,6\}})*\mathbf{K}) = \mathbf{u}_0\cdot(\mathbf{e}_{\{2,4,6\}}*\mathbf{K}) \oplus \mathbf{u}_0\cdot\mathbf{X}\oplus\mathbf{v}\cdot\mathbf{Y}$$

or

$$\mathbf{u}_0\cdot((1\oplus \mathbf{e}_{\{2,4,6\}})*\mathbf{K}) = \mathbf{u}_0\cdot(\mathbf{e}_{\{2,4,6\}}*\mathbf{K}) \oplus \mathbf{u}_0\cdot\mathbf{X}\oplus\mathbf{v}\cdot\mathbf{Y}\oplus 1$$

is true with probability 0.994, depending on whether $W_{F,\{2,4,6\}}^{(0.99)}(\mathbf{u}_0, \mathbf{e}_1)$ is positive or negative, respectively.

In Table 1, we list $\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{v}\cdot F)$ and $\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\mathcal{S}}^{(p)}(\mathbf{u}, \mathbf{v}\cdot F)$ for all 8 DES S-boxes.

F	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{v}\cdot F)$	0.219	0.219	0.219	0.156	0.219	0.188	0.281	0.188
$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon_{\mathcal{S}}^{(p)}(\mathbf{u}, \mathbf{v}\cdot F)$	0.494	0.494	0.497	0.489	0.494	0.491	0.494	0.494

TABLE 1. Maximum bias without and with biased inputs for all DES S-boxes.

4.2. EXPERIMENTS WITH AES S-BOX. Let F be the AES S-box. The truth table of $\mathbf{e}_1 \cdot F$ is

$$(0000100100001110111010011111110111100011011101000101010100111010010000010011010101010110011000011111000101000011010111011110011110010100110100110100110000000010100111011010011000000011101110110011010011001010000111110001110010001000011101111001111011100111101100010010100)$$

From the direct computation we observe that

$$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon(\mathbf{u}, \mathbf{e}_1 \cdot F) \approx 0.063,$$

whereas

$$\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon^{(0.99)}_{\{2,3,5\}}(\mathbf{u}, \mathbf{e}_1 \cdot F) \approx 0.334. \tag{13}$$

Suppose that $\max_{\mathbf{u}\in\mathbb{F}_2^n} \epsilon^{(0.99)}_{\{2,3,5\}}(\mathbf{u}, \mathbf{e}_1 \cdot F) = \epsilon^{(0.99)}_{\{2,3,5\}}(\mathbf{u}_0, \mathbf{e}_1 \cdot F)$. Therefore,

$$\Pr[\mathbf{u}_0 \cdot ((1 \oplus \mathbf{e}_{\{2,3,5\}}) * \mathbf{K}) = \mathbf{u}_0 \cdot (\mathbf{e}_{\{2,3,5\}} * \mathbf{K}) \oplus \mathbf{u}_0 \cdot \mathbf{X} \oplus \mathbf{v} \cdot \mathbf{Y}] \in \{0.834, 0.166\}.$$

It seems that if we assume values of $\mathbf{e}_{\{2,3,5\}} * \mathbf{K}$ and choose $\mathbf{X}$ as described in the attack model, then either

$$\mathbf{u}_0 \cdot ((1 \oplus \mathbf{e}_{\{2,3,5\}}) * \mathbf{K}) = \mathbf{u}_0 \cdot (\mathbf{e}_{\{2,3,5\}} * \mathbf{K}) \oplus \mathbf{u}_0 \cdot \mathbf{X} \oplus \mathbf{v} \cdot \mathbf{Y}$$

or

$$\mathbf{u}_0 \cdot ((1 \oplus \mathbf{e}_{\{2,3,5\}}) * \mathbf{K}) = \mathbf{u}_0 \cdot (\mathbf{e}_{\{2,3,5\}} * \mathbf{K}) \oplus \mathbf{u}_0 \cdot \mathbf{X} \oplus \mathbf{v} \cdot \mathbf{Y} \oplus 1$$

is true with probability 0.834, depending on whether $W^{(0.99)}_{F,\{2,3,5\}}(\mathbf{u}_0, \mathbf{e}_1)$ is positive or negative, respectively.

5. CONCLUSION

Linear cryptanalysis of block ciphers involving S-boxes requires approximation of component functions of some of the S-boxes by affine functions. Typically, linear cryptanalysis assumes that the inputs to the S-boxes are uniformly randomly distributed over the set of all binary strings of the same length as the data-width. Analysis of the S-boxes, in case the input distribution is biased, has remained an open problem so far. In this paper, for the first time we generalized the concept of non-linearity of S-boxes with biased inputs. We showed that the typical case of uniform distribution is a special case of our generalized analysis. Moreover, we outline a chosen-plaintext attack model that can exploit the above analysis.

Our results establish a new design criteria for cryptographically secure S-boxes. More research is needed in this direction to explore the possibility of efficient practical cryptanalysis using our approach.

REFERENCES

[1] P. Erdös and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960), 17–61.

[2] E. Friedgut and Gil Kalai, *Every monotone graph property has a sharp threshold*, Proc. AMS 124 (10) (1996), 2293–3002.

[3] S. Gangopadhyay, A. Kar Gangopadhyay, S. Pollatos and P. Stănică, *Cryptographic Boolean functions with biased inputs*, Cryptography and Communications - Discrete Structures and Sequences 9:2 (2017), 301–314.

[4] Y. Lu and Y. Desmedt, *Bias analysis of a certain problem with applications to E0 and Shannon ciper*, ICISC 2010, LNCS 6829, 2011, pp. 16–28.

[5] M. Matsui, *Linear cryptanalysis method for DES cipher*, EUROCRYPT'93, LNCS 765, (Springer) 1994, 386–397.
[6] R. O'Donnell, Analysis of Boolean functions, Cambridge University Press, 2014.
[7] M. G. Parker, *Generalised S-box nonlinearity*, NESSIE Public Document, 11.02.03: NES/DOC/UIB/WP5/020/A.
[8] D. R. Stinson. *Cryptography: Theory and Practice*, Third Edition, Chapman and Hall/CRC, 2005.