Spectral Characterization and Analysis of Avalanche in Cryptographic Substitution Boxes using Walsh-Hadamard Transformations

Fouz Sattar
Dept. of Electrical Engineering
University of Engineering and
Technology, Taxila, Pakistan

Muid Mufti
Dept. of Electrical Engineering
University of Engineering and
Technology, Taxila,Pakistan

ABSTRACT
This paper presents a novel analytical technique of examining the avalanche characteristics of cryptographic substitution boxes (s-boxes) using the Walsh Hadamard spectral analysis of their Boolean maps. Avalanche is a desirable cryptographic property that is necessary to ensure that a cipher is not susceptible to statistical attacks and small difference between two plaintexts results in a significant random difference between the two corresponding ciphertexts. An analytical model of avalanche characteristics based on spectral analysis is developed which allows us to determine the degree of avalanche achievable in an $n \times m$ s-box. Analytical results are finally used to investigate the avalanche characteristics of the AES s-box.

Keywords
Avalanche criteria, Strict Avalanche Criteria, S-box, Boolean function, Walsh-Hadamard transform, Cryptography

1. INTRODUCTION
Symmetric key block ciphers based on the Feistel substitution permutation network (SPN) [1] generally utilize substitution boxes (s-boxes) as a means to introduce nonlinearly to the overall cryptographic algorithm.

A $n \times m$ s-box can be viewed as a mapping function, which maps n-bit input vectors nonlinearly into m-bit output vectors. To yield a cryptographically strong SPN, the underlying s-boxes should satisfy certain dynamic properties [2] that guarantee the randomness of SPN. Avalanche [3] is one of such cryptographic properties which states that a small number of bit differences in the input of a function results in ``avalanche'' of changes in subsequent round functions resulting in large number of output bit changes. More formally, the concept of avalanche in a function can be defined as the property of one bit change in the input leading to a change in, on average, half of the output bits [4].

An extension to this definition was proposed by Webster and Tavares [5] and is referred to as Strict Avalanche Criteria (SAC). A function satisfies SAC if each output bit changes with probability ½ when a single input bit is complemented.

Avalanche and Strict Avalanche Criteria ensure that the cryptosystem does not exhibit any statistical correlation between input and output that a cryptanalyst might use in an attack. For example, if some output bits of a cryptosystem depend only on the few input bits, then, by observing a significant number of input output pairs, such as in chosen plaintext attack, a cryptanalyst might be able to detect these relations and use this information to aid in the exhaustive key search. However, it should also be noted that satisfaction of these criteria is not sufficient to ensure the security of the cryptosystem and cryptanalysis techniques such as differential and linear cryptanalysis [9] have been applied effectively on on ciphers which have been shown to reasonably satisfy SAC.

Study of Avalanche and SAC characteristics of cryptosystems has been the area of focus for several authors [2,4,6,10,15,16,17,19,20]. Heys and Tavares developed stochastic and deterministic permutation frameworks [4] to model avalanche in SPNs. Their work was extended in [16] to develop a model of avalanche characteristics of ciphers similar in structure to Data Encrpytion Standard (DES). Yücel and Vergili [2,6] have experimentally studied the statistics of Avalanche properties over ensembles of randomly chosen s-boxes.

The results presented in this paper are a contribution towards development of an analytical model to quantify the avalanche characteristics of s-boxes using spectral techniques based on Walsh-Hadamard transform. Such techniques have been found useful in Boolean function classification [8] and differential and linear cryptanalysis. Conditions on Walsh spectra of a Boolean function guaranteeing satisfaction of SAC were first examined by Forré [9]. This paper extends Forre formalism to characterize and analyze avalanche in an $n \times m$ s-box by considering it as a vector output Boolean function. A relationship between Avalanche probability and Walsh spectrum of s-box is established which can determine the degree of avalanche achievable in practice.

The rest of this paper is organized as follows: In Section 2 we introduce the main notations, preliminaries and definitions used throughout the paper. In Section 3 we present Walsh Hadamard transform based spectral techniques and extension of Forré formalism to develop an analytical model for the avalanche characteristics of s-boxes. In Section 4, we use the analytical results to investigate the avalanche characteristics of AES s-box. Finally in Section 5, we present some concluding remarks.

2. PRELIMINARIES

2.1 Avalanche and Strict Avalanche Criteria

Let c denote a $n \times m$ mapping function and p and p_i denote two input vectors differing only in bit $i \ \forall \ 1 \le i \le n$ such that $d^{(i)} = c(p) \oplus c(p_i) = [d_1^{(i)} d_2^{(i)} ... d_m^{(i)}]$, $d_k^{(i)} \in Z_2 \ \forall \ 1 \le k \le m$ and $W_k^{(i)} = \sum_{p \in Z_2^n} d_k^{(i)}$

Definition 2.1.1 [6,7]:

The function c is said to satisfy the avalanche criterion if:

$$\frac{\sum_{k=1}^{m} W_k^{(i)}}{m \cdot 2^n} = \frac{1}{2} \quad \forall \quad 1 \le i \le n$$

Definition 2.1.2:

The function c is said to satisfy the strict avalanche criterion if:

$$\frac{W_k^{(i)}}{2^n} = \frac{1}{2} \quad \forall \quad 1 \le i \le n \quad and \quad 1 \le k \le m$$

2.2 Boolean Functions and Walsh-Hadamard Transform

Let $f(x): Z_2^n \to Z_2$ be a Boolean function with domain Z_2^n that takes values 0 and 1. Z_2^n denotes the n-dimensional vector space of binary n-tuples $(x_1, x_2 ... x_n)$ over finite field $GF(2)$ with modulo-2 addition $\oplus$.

The **Walsh-Hadamard transform** of $f(x)$ is the real-valued function over the vector space Z_n^2 defined as:

$$\Omega(f(x)) = F(w) = \sum_{x \in Z_2^n} f(x) \cdot (-1)^{x \cdot w} \qquad (1)$$

where $w \in Z_2^n$ and $x \cdot w$ denotes canonical dot-product of x and w in Z_n^2, defined as:

$$x \cdot w = x_1 w_1 \oplus x_2 w_2 \ldots \oplus x_n w_n$$

The function $f(x)$ can be recovered by the **inverse Walsh-Hadamard transform**:

$$\Omega^{-1}(F(w)) = f(x) = \frac{1}{2^n} \sum_{w \in Z_2^n} F(w) \cdot (-1)^{x \cdot w} \qquad (2)$$

3. S-BOX AVALANCHE COMPUTATION

Consider an $n \times m$ s-box with n-bit input and m-bit output, represented as a vector mapping $s(x): Z_2^n \to Z_2^m$. We decompose this mapping into m-component Boolean functions $(s_1(x), s_2(x) \ldots s_m(x))$ where $s_j(x): Z_2^n \to Z_2 \ \forall \ j = 1 \ldots m$.

Considering the j-th component Boolean function $s_j(x)$, let x and x_i denote two n-bit vectors such that x and x_i differ only in bit $i \ \forall \ 1 \le i \le n$ and let $a_j^{(i)}$ denote the corresponding output difference i.e.

$$a_j^{(i)} = s_j(x) \oplus s_j(x_i) \qquad (3)$$

If c_i represents the n-dimensional unit vector with a one at the i-th place and zeros elsewhere, then (3) can alternatively be written as:

$$a_j^{(i)} = s_j(x) \oplus s_j(x \oplus c_i) \qquad (4)$$

If p_i number of n-tuples $x \in Z_2^n$ satisfy $s_j(x) = s_j(x \oplus c_i)$ and q_i number of n-tuples $x \in Z_2^n$ satisfy $s_j(x) \ne s_j(x \oplus c_i)$ then the avalanche probability $P_j^{(i)}$ of the j-th component Boolean function $s_j(x)$ can be written as:

$$P_j^{(i)} = \frac{q_i}{2^n} \qquad (5)$$

$P_j^{(i)}$ can be interpreted as the probability of change of the output of $s_j(x)$ when only the i-th bit of the input $x \in Z_2^n$ is changed. Let $\hat{s}_j(x)$ denote a function which converts the range $\{0,1\}$ of $s_j(x)$ to range $\{1,-1\}$ i.e.

$$\hat{s}_j(x) = 1 - 2 s_j(x) \qquad (6)$$

$$\hat{a}_j^{(i)} = \hat{s}_j(x) \cdot \hat{s}_j(x \oplus c_i) \qquad (7)$$

then $\hat{a}_j^{(i)} = +1$ for p_i of n-tuples $x \in Z_2^n$ and $\hat{a}_j^{(i)} = -1$ for q_i of n-tuples $x \in Z_2^n$.

The summation of $\hat{a}_j^{(i)}$ over all $x \in Z_2^n$ can therefore be expressed as:

$$\sum_{x \in Z_2^n} \hat{a}_j^{(i)} = \sum_{x \in Z_2^n} (\hat{s}_j(x) \cdot \hat{s}_j(x \oplus c_i)) = p_i - q_i \tag{8}$$

The summation $\sum_{x \in Z_2^n} (\hat{s}_j(x) \cdot \hat{s}_j(x \oplus c_i)) = \zeta$ can be written as:

$$\begin{aligned}
\zeta &= \Omega^{-1}[\sum_{c_i \in Z_2^n} \sum_{x \in Z_2^n} \hat{s}_j(x) \cdot \hat{s}_j(x \oplus c_i)(-1)^{c_i \cdot w}] \\
&= \Omega^{-1}[\sum_{x \in Z_2^n} \hat{s}_j(x)(\sum_{c_i \in Z_2^n} \hat{s}_j(x \oplus c_i)(-1)^{c_i \cdot w})] \\
&= \Omega^{-1}[\sum_{x \in Z_2^n} \hat{s}_j(x)(\sum_{v \in Z_2^n} \hat{s}_j(v)(-1)^{(x \oplus v) \cdot w})] \\
&= \Omega^{-1}[\sum_{x \in Z_2^n} \hat{s}_j(x)(\sum_{v \in Z_2^n} \hat{s}_j(v)(-1)^{v \cdot w})(-1)^{x \cdot w}] \\
&= \Omega^{-1}[(\sum_{x \in Z_2^n} \hat{s}_j(x)(-1)^{x \cdot w})\hat{S}_j(w)] \\
&= \Omega^{-1}[\hat{S}_j(w)\hat{S}_j(w)] \\
&= p_i - q_i
\end{aligned} \tag{9}$$

Solution of (9) for q_i and substitution in (5) yields:

$$\begin{aligned}
P_j^{(i)} &= \frac{1}{2} - 2^{-(n+1)}[\Omega^{-1}(\hat{S}_j(w) \cdot \hat{S}_j(w))](c_i) \\
&= \frac{1}{2} - 2^{-(2n+1)} \sum_{w \in Z_2^n} (-1)^{c_i \cdot w} \hat{S}_j^2(w) \\
&= \frac{1}{2} - 2^{-(2n+1)} \sum_{w \in Z_2^n} (-1)^{w_i} \hat{S}_j^2(w)
\end{aligned} \tag{10}$$

Finally, the probability of change of the overall output bits of the s-box when the i-th input bit is complemented is given as:

$$\begin{aligned}
P_{aval}^{(i)} &= \frac{\sum_{j \in \{1 \ldots m\}} P_j^{(i)}}{m} \\
&= \frac{1}{2} - \frac{2^{(2n+1)}}{m} \sum_{j \in \{1 \ldots m\}} \sum_{w \in Z_2^n} (-1)^{w_i} \hat{S}_j^2(w)
\end{aligned} \tag{11}$$

The relationship between $S_j(w)$ and $\hat{S}_j(w)$ is given by [5]:

$$\hat{S}_j(w) = -2S_j(w) + 2^n \delta(w) \tag{12}$$

where δ denotes the Kronecker delta defined as:

$$\delta = \begin{cases} 1, \forall \quad w = 0 \\ 0, else \end{cases}$$

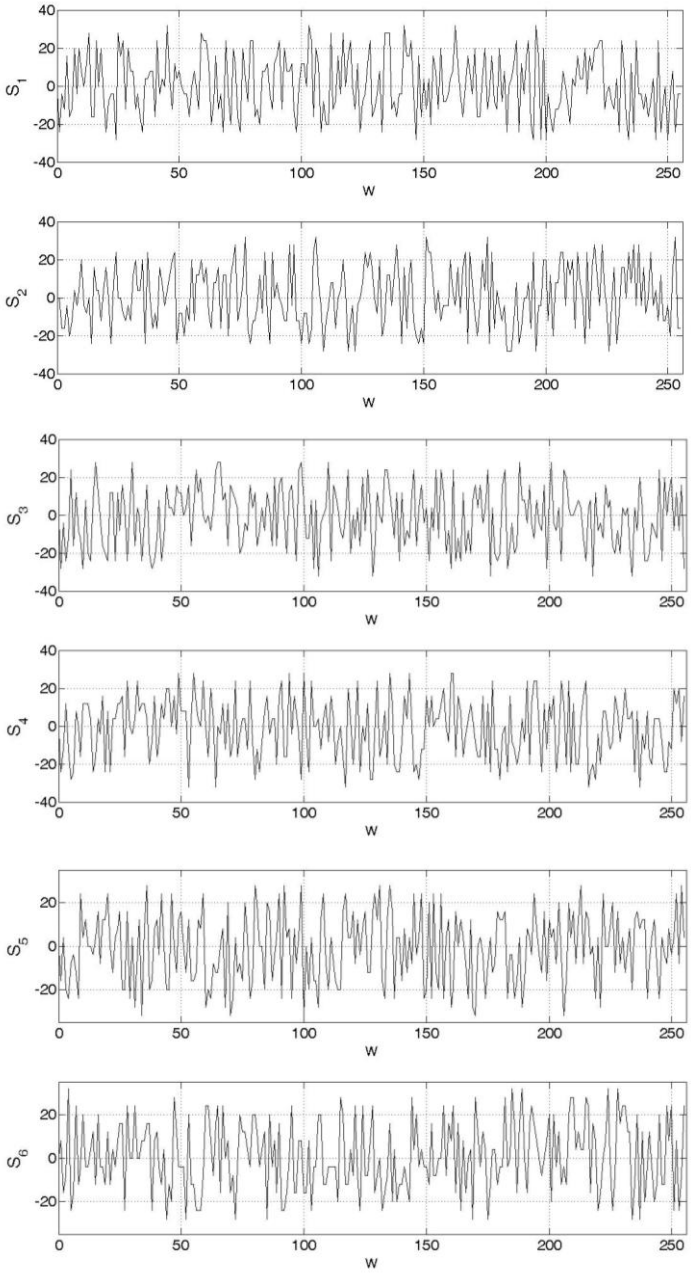
4. AES S-BOX ANALYSIS

In this section, we present the results of the computations of the preceding section for Advanced Encryption Standard (AES) s-box [18]. The 8×8 AES s-box is based on the mapping $x \to x^{-1}$ followed by an affine transformation over $GF(2)$, where x^{-1} denotes the multiplicative inverse in $GF(2^8)$.

One of the design criteria for the AES s-box is resistance against diffrential and linear cryptanalysis by minimization of the largest non-trivial correlation between linear combinations of input bits and linear combination of output bits [13,14,12], which in turn results in good Avalanche characteristics.

The AES s-box $s(x)$ can be decomposed into eight component Boolean functions $(s_1(x), s_2(x) \ldots s_8(x))$ specified as $s_k(x) = s(x) \& (1 << (k-1)) \quad \forall \quad k = 1 \ldots 8$, so that

$s_1(x)$ therefore represents the function taken from the least significant bit of each s-box value and $s_8(x)$ represents the function taken from the most significant bit of each s-box value.The Walsh spectra of the eigth component functions are shown in Figure 1.



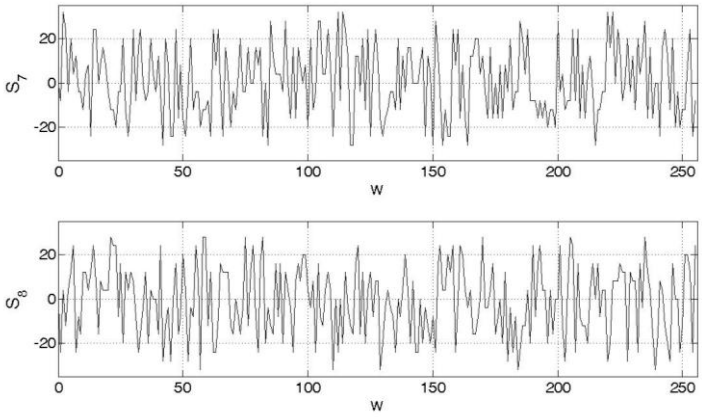


Fig 1: Walsh Spectra of AES s-box component functions

The Avalanche probabilities are finally computed using the spectral values and (10),(11) and (12). The results are summarized in Table 1. It is clear that AES s-box does not exactly satisfy the Avalanche and Strict Avalanche criteria but the Avalanche probability $P_{aval}^{(i)}$ is found to be close to 0.5.

The relative errors ε_{Aval} and ε_{SAC} for the Avalanche and Strict Avalanche criteria respectively are defined as [2]:

$$\varepsilon_{Aval} = max|2P_{aval}^{(i)} - 1| \quad \forall \quad 1 \le i \le n \tag{13}$$

$$\varepsilon_{SAC} = max|2P_{j}^{(i)} - 1| \quad \forall \quad 1 \le i \le n \quad and \quad 1 \le j \le m \tag{14}$$

From Table 1, the relative errors ε_{Aval} and ε_{SAC} are found to be 0.0352 and 0.1250 respectively which are in agreement with the experimental results reported in [7].

Table 1: Avalanche Probabilities of AES s-box

i	$P_8^{(i)}$	$P_7^{(i)}$	$P_6^{(i)}$	$P_5^{(i)}$	$P_4^{(i)}$	$P_3^{(i)}$	$P_2^{(i)}$	$P_1^{(i)}$	$P_{aval}^{(i)}$
1	0.5000	0.4531	0.4844	0.4531	0.5625	0.4531	0.5156	0.5156	0.4922
2	0.5313	0.5000	0.4531	0.4844	0.5000	0.5625	0.4844	0.4688	0.4980
3	0.5000	0.5313	0.5000	0.5625	0.4688	0.5000	0.5156	0.5156	0.5117
4	0.5469	0.5000	0.5313	0.5000	0.4531	0.4688	0.5313	0.5313	0.5078
5	0.5313	0.5469	0.5000	0.5000	0.5156	0.4531	0.5000	0.4531	0.5000
6	0.5313	0.5313	0.5469	0.4688	0.4688	0.5156	0.5156	0.4531	0.5039
7	0.4844	0.5313	0.5313	0.4688	0.5156	0.4688	0.5313	0.5313	0.5078
8	0.5156	0.4844	0.5313	0.4844	0.5313	0.5156	0.5625	0.5156	0.5176

5. CONCLUSIONS

We have presented a spectral analysis technique to model the avalanche characteristics of substitution boxes. It has been shown that the degree of Avalanche practically achievable in s-boxes can be characterized in Walsh domain. The spectral model is applied on the AES s-box and results are found to be in agreement with the previously reported experimental results.

The model can used to validate the Avalanche properties of cryptographically secure s-boxes.

Extension of the analysis to investigate other dynamic properties such as bit independence, correlation of avalanche vectors etc. could form an important direction for future study.

6. REFERENCES

[1] H. Feistel, Cryptography and computer privacy, Scientific American, vol. 228, no. 5, pp. 15--23, 1973.

[2] Isil Vergili and Melek D.Yücel, On Satisfaction of Some Security Criteria for Randomly Chosen S-Boxes, Proc. 20th Biennial Symp. on Communications, pp.64-68, Kingston, Ontario, Canada, May 2000.

[3] H. Feistel, W. Notz, and J. L. Smith, Some Cryptographic Techniques for Machine-to-Machine Data Communications, Proceedings of the IEEE, 63 (1975), pp. 1545-1554.

[4] H.M. Heys and S.E. Tavares, Avalanche Characteristics of Substitution-Permutation Encryption Networks, IEEE Trans. on Computers, vol. 44, no. 9 , pp. 1131-1139, 1995.

[5] A. F. Webster and S. E. Tavares, On the design of S-boxes, Advances in Cryptology: Proceedings of CRYPTO '85, Springer-Verlag, Berlin, pp. 523--534, 1986.

[6] Isil Vergili and Melek D. Yücel, Avalanche and Bit Independence Properties for the Ensembles of Randomly Chosen nxn S-Boxes, Turkish Journal of Electrical Eng. and Computer Sciences, Vol.9, No.2, pp.137-145, August 2001.

[7] Selçuk Kavut and Melek D.Yücel, On Some Cryptographic Properties of Rijndael, Lecture Notes in Computer Science: Information Assurance in Computer Networks, Methods, Models and Architectures for Network Security, LNCS Vol.2052, Springer-Verlag, pp.300-311, May 2001.

[8] A. Bernasconi, B. Codenotti, Spectral Analysis of Boolean Functions as a Graph Eigenvalue Problem, IEEE Transactions on Computers, Vol. 48(3) (1999), 345-351.

[9] Joan Daemen, Cipher and hash function design strategies based on linear and differential cryptanalysis, Doctoral dissertation, KU Leuven 1995.

[10] Réjane Forré, The strict avalanche criterion: Spectral properties of Boolean functions and an extended definition, Advances in Cryptology - CRYPTO '88, 450-468.

[11] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts, and J. Vandewalle, Propagation characteristics of boolean functions, Advances in Cryptology: Proceedings of EUROCRYPT '90, Springer-Verlag, Berlin, 1991, pp. 161-173.

[12] K.Nyberg, Differentially uniform mappings for cryptography, Advances in Cryptology, Proceedings of Eurocrypt'93, LNCS 765, T.Helleseth, Ed.,Springer-Verlag, 1994, pp.55-64.

[13] Daemen and V.Rijmen, AES Proposal: Rijndael, Proceedings of First Advanced Encryption Standard (AES) Conference, California, August 1998.

[14] Daemen, Annex to AES Proposal Rijndael, Chapter 5 Propagation and Correlation, June 1998.

[15] S. Mister, C. M. Adams, Practical S-Box Design, SAC'96 Third Annual Workshop on Selected Areas in Cryptography, Queen's University, Kingston, pp. 61-76, 1996.

[16] H.M. Heys, Avalanche Characteristics of DES-like Ciphers, Proceedings of SAC '96 - Workshop on Selected Areas in Cryptography, Queen's University, Kingston, Ontario, Aug. 1996.

[17] H.M. Heys and S.E. Tavares, Key Clustering in Substitution-Permutation Network Cryptosystems, Proceedings of SAC '94 - Workshop on Selected Area in Cryptography, Kingston, Ontario, May 1994.

[18] National Institute of Standards and Technology. FIPS Pub 197: Advanced Encryption Standard (AES), November 2001

[19] Kaoru Kurosawa and Takashi Satoh, Generalization of higher order SAC to vector output Boolean function, IEICE Trans. E81-A, pp.41-47 (1998).

[20] Kwangjo Kim, Construction of DES-like S-boxes Based on Boolean Functions Satisfying the SAC, ASIACRYPT 1991: 59-72