2019 32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems (VLSID)

A Machine Learning Based Approach to Predict Power Efficiency of S-boxes

Rajat Sadhukhan
Indian Institute of Technology-Kharagpur
Kharagpur, West Bengal, India
Email: rajat.sadhukhan@iitkgp.ac.in

Nilanjan Datta
Indian Institute of Technology-Kharagpur
Kharagpur, West Bengal, India
Email: nilanjan_isi_jrf@yahoo.com

Debdeep Mukhopadhyay
Indian Institute of Technology-Kharagpur
Kharagpur, West Bengal, India
Email: debdeep.mukhopadhyay@gmail.com

***Abstract*—In the era of lightweight cryptography, designing cryptographically good and power efficient 4×4 S-boxes is a widely discussed problem. While the optimal cryptographic properties are easy to verify, it is not very straightforward to verify whether a S-box is power efficient or not. The traditional approach is to explicitly determine the dynamic power consumption using commercially available CAD tools and report accordingly based on a pre-defined threshold value. However, this procedure is highly time consuming, and the overhead becomes formidable while dealing with a set of S-boxes from a large space. This mandates development of an automation tool which should be able to quickly characterize the power efficiency from the Boolean function representation of an S-box. In this paper, we present a supervised machine learning (ML) assisted automated framework to resolve the problem for 4×4 S-boxes, which turns out to be approximately 14 times faster (using AND-OR-NOT gates) than the traditional approach. The key idea is to extrapolate the knowledge of the literal counts of various functional forms, AND-OR-NOT gate counts in the simplified SOP form of the underlying Boolean functions corresponding to the S-box to predict the dynamic power efficiency. We demonstrate the effectiveness of our framework by reporting a set of power efficient S-boxes from a large set of 4×4 optimal S-boxes. The experimental results and performance of our novel technique depicts its superiority with high efficiency and low time overhead.**

***Index Terms*—Power Efficiency, Optimal S-box, Dynamic power, Machine Learning**

I. Introduction

The heavy resource constraints on IoT devices make it impossible to run conventional cryptographic algorithms, which lead to the development of lightweight cryptographic algorithms and primitives. Recently, this development of lightweight cryptographic primitives has gained its momentum with the announcement of lightweight cryptographic project by NIST [1]. As S-boxes are the basic building blocks (used to provide the non-linearity) for designing block ciphers, designing cryptographically good and power efficient S-boxes is a widely discussed problem.

A. Determining Power Efficiency of S-boxes: The Traditional and Our Approach

An illustration of traditional approaches(Probabilistic Approach [2],Simulation Based Technique [3]), along with our proposed approach is shown in Fig.1.

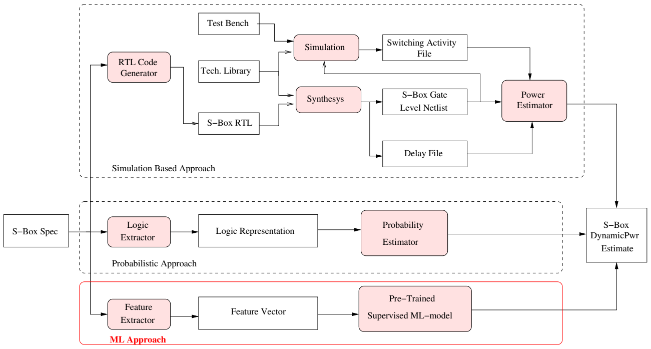


Fig. 1. Design Flow

B. Our Contribution

Central to this work is power efficiency of S-boxes and contributions are two folded. First, we present a supervised machine learning assisted automated framework to classify a set of $n \times n$ S-boxes into two classes based on their power efficiency. Next, we demonstrate the effectiveness of our framework by reporting a set of power efficient S-boxes from a large set of 4×4 optimal S-boxes. The experimental results shows that our tool is approximately 84% accurate for both the classifiers.

C. Significance of the Work

To the best of our knowledge, this is the first machine learning based approach that predicts whether an S-box is power efficient or not. Our algorithm requires roughly 0.874

2380-6923/19/$31.00 ©2019 IEEE
DOI 10.1109/VLSID.2019.00121

seconds to determine power efficiency of an S-box, in an Intel Xeon machine, operating at 2 Ghz. We have also used Synopsys *Design Compiler* to compute the power of an S-box, which takes roughly 11.843 seconds in the same machine. Our algorithm reduces the time overhead by a factor of around 14 times.

II. Proposed Methodology

We have followed supervised learning approach to construct a binary classifier for our problem. Given a predefined threshold power P_{th}, the classification is done as follows:

(i) $S \in$ Class 0 (bad), if $P_{\text{dyn}}(S) > P_{\text{th}}$
(ii) $S \in$ Class 1 (good), if $P_{\text{dyn}}(S) \leq P_{\text{th}}$.

We have mapped our problem to classification rather than regression as we are interested to exploit large search space and report whether a set of S-boxes are power efficient or not, rather than reporting the exact dynamic power of an S-box. The formal algorithm corresponding our framework is presented in Algorithm 1.

Algorithm 1: ML-Framework for Classifying S-boxes

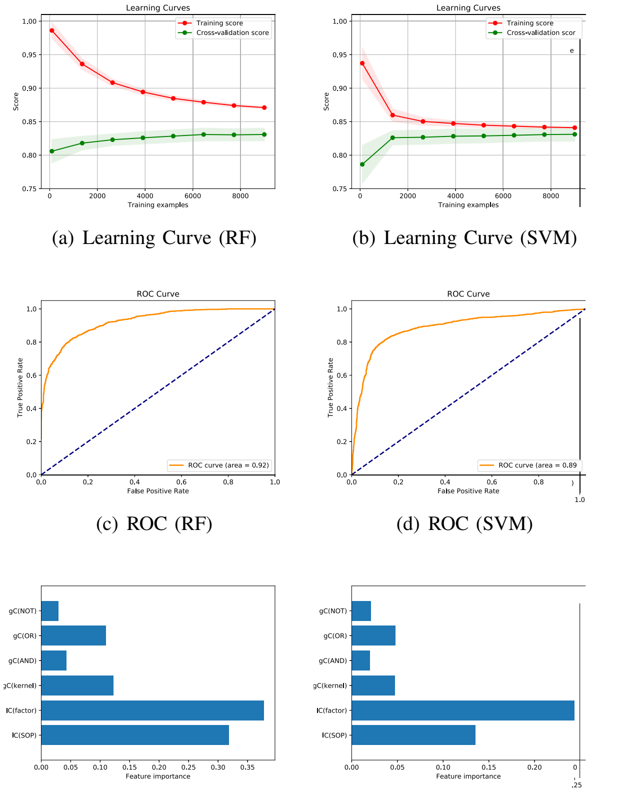
```
Input : $\Sigma_n \subset \mathsf{Perm}(2^n)$, a set of S-boxes having
        property $P$.
Output: Classification vector $E_{\text{test}}$ of S-boxes
        corresponding to the test vector $T_{\text{test}}$
Construct $T_{\text{train}} \subset \Sigma_n$ and $T_{\text{test}} = \Sigma_n \backslash T_{\text{train}}$ ;
Let $T_{\text{train}} = (S_1, \ldots, S_t)$ and $T_{\text{test}} = (S_{t+1}, \ldots, S_{|\Sigma_n|})$
for $i = 1;\ i \leq t;\ i = i + 1$ do
    $F_{S_i} = \mathsf{Extract\_Feature}(S_i)$;
    $E_{S_i} = \mathsf{Find\_Class}(S_i)$;
end
/* Train Model */
$C_{\text{train}}$ =TrainModel$((F_{S_1}, \ldots, F_{S_t}),\ (E_{S_1}, \ldots, E_{S_t}))$
/* Classify*/
for $i = t + 1;\ i \leq |\Sigma_n|;\ i = i + 1$ do
    $F_{S_i} = \mathsf{Extract\_Feature}(S_i)$;
    $E_{S_i} = \mathsf{Predict\_Class}(F_{S_i}, C_{\text{train}})$;
end
$E_{\text{test}} = (E_{t+1}, \ldots, E_{|\Sigma_n|})$;
return $E_{\text{test}}$;
}
```

III. Application to 4 X 4 Optimal S-box

A. Setup

In order to evaluate our proposed methodology we started with a list of 10000 cryptographically strong S-boxes having differential uniformity and non-linearity of 4 obtained using genetic algorithm. To extract features of every S-box we used *Sequential Interactive System(SIS)*version 1.3 and *Espresso*version 2.3.In the synthesis process we used standard cell library ($180nm$) consisting of only 2-input AND, OR, and NOT gate from TSL18FS120 cell library. Finally, to build our ML model and data analysis we used *Scikit-Learn*ML tool (v0.19).

B. Performance



(a) Learning Curve (RF)

(b) Learning Curve (SVM)

(c) ROC (RF)

(d) ROC (SVM)

(e) Feature Importance (RF)

(f) Feature Importance (SVM)

Fig. 2. Learning, ROC Curves and Feature Importance Chart corresponding to RF and SVM Model

Classifier	TN	FP	FN	TP	Accuracy	F1 Score
RF	1296	197	274	1233	84.3%	0.846
SVM	1310	203	296	1191	83.4%	0.840

TABLE I
Performance Result for 4 × 4 Optimal S-boxes

IV. Conclusion and Future Works

In this paper, we have presented a supervised machine learning assisted automated framework to report the power efficiency of cryptographically strong S-boxes. Extension of this ML based methodology to the regression problem of predicting the dynamic power value of an S-box and thereby reporting the best S-box from a set of S-boxes seems to be an intriguing future direction.

References

[1] K. A. McKay, L. Bassham, M. S. Turan, and N. Mouha, “Report on Lightweight Cryptography,” 2017.

[2] F. N. Najm, “A survey of power estimation techniques in vlsi circuits,” *IEEE Transactions on VLSI Systems*, vol. 2, no. 4, pp. 446–455, 1994.

[3] S. Picek, B. Yang, V. Rozic, and N. Mentens, “On the construction of hardware-friendly 4x4 and 5x5 s-boxes,” in *SAC 2016*.