Deep Learning Based Side Channel Attacks on Lightweight Cryptography (Student Abstract)

Alexander Benjamin [1], **Jack Herzoff** [2], **Liljana Babinkostova** [2] and **Edoardo Serra** [2]
[1]Brown University
[2]Boise State University
alexander_benjamin@brown.edu, jackherzoff@u.boisestate.edu, liljanababinkostova @boisestate.edu, edoardoserra@boisestate.edu

Abstract

Computing devices continue to be increasingly spread out within our everyday environments. Computers are embedded into everyday devices in order to serve the functionality of electronic components or to enable new services in their own right. Existing Substitution-Permutation Network (SPN) ciphers, such as the Advanced Encryption Standard (AES), are not suitable for devices where memory, power consumption or processing power is limited. Lightweight SPN ciphers, such as GIFT-128 provide a solution for running cryptography on low resource devices. The GIFT-128 cryptographic scheme is a building block for GIFT-COFB (Authenticated Encryption with Associated Data), one of the finalists in the ongoing NIST lightweight cryptography standardization process (NISTIR 8369). Determination of an adequate level of security and providing subsequent mechanisms to achieve it, is one of the most pressing problems regarding embedded computing devices. In this paper we present experimental results and comparative study of Deep Learning (DL) based Side Channel Attacks on lightweight GIFT-128. To our knowledge, this is the first study of the security of GIFT-128 against DL-based SCA attacks.

Introduction

Computing devices continue to be increasingly spread out within our everyday environments. Computers are "embedded" into everyday devices in order to serve the functionality of electronic components or to enable new services in their own right. Determination of an adequate level of security and providing subsequent mechanisms to achieve it, is one of the most pressing problems regarding embedded computing devices.The fundamental problem for resource-constrained systems is the fact that current cryptographic algorithms (e.g. (Daemen and Rijmen 2003)) utilize significant energy consumption and storage overhead. In 2016, the National Institute of Standards and Technology (NIST) launched a multi-year standardization process for lightweight cryptography (LWC), in which ciphers are evaluated for efficient implementation on resource-constrained platforms and sufficient level of security (Arribas 2020). Side-channel analysis (SCA)is a significant threat to the successful deployment of cryptographic solutions. Previous research at the intersection between Deep Learning (DL) and SCA focused (almost) exclusively on the AES (Daemen and Rijmen 2003). Thus, a consistent study of the SCA-vulnerabilities of the NIST LWE candidates for a standard, is still missing. In this paper we investigate and provide experimental results of DL-based SCA on GIFT-128 (Banik et al. 2017), the building block of one of the finalists in the ongoing NIST LWE standardization process.

Background

In this section, we introduce general cryptographic terminology and present a background on side channel attacks.

An **iterated block cipher** is a block cipher obtained by iterating r times a round function $R : \{0,1\}^n \to \{0,1\}^n$, each time with its own key $K_i \in \mathcal{K}$, where $\mathcal{K}$ is called round key space. The block size is n bits, the number of rounds is equal to r, $X^{(0)}$ is the plaintext, and $X^{(r)}$ is the ciphertext $X^{(i)} = R_{K_i}(X^{(i-1)})$ for $1 \leq i \leq r$.

Side-Channel Attacks (SCA) analyze physical leakage that is emitted during the execution of a cryptographic algorithm in a device (e.g. power consumption (P. Kocher, J. Jaffe, and B. Jun 1999)). Deep Learning (DL) techiques have shown to be effective in various scenarios (e.g. (Timon 2019), (E. Prouff, R. Strullu, R. Benadjila, E. Cagli, and C. Dumas 2018)). While the use of DL methods in so-called profiled SCA have been well studied, a DL-based non-profiling SCA techniques have been proposed only recently (Timon 2019) and applied only to conventional cryptographic standards (Daemen and Rijmen 2003).

Correlation Power Analysis (CPA) is a type of SCA described as follows: Given a set of power traces and the corresponding sets of intermediate values $\phi_1, \phi_2, \ldots \phi_{2^{|k|}}$, CPA aims at recovering the secret subkey k^* using a correlation factor between the measured power samples and the power model of the computed sensitive values (E. Brier, C. Clavier, and F. Olivier 2004). A power model is used to determine the hypothetical power consumption of the target device as a function of the intermediate value ϕ_k considering the power consumption of the device.

Experimental Results

We perform experiments using power traces of GIFT-128 (Banik et al. 2017) collected from the ChipWhisperer-Lite board (W. Unger, L. Babinkostova, M. Borowczak and R.

Erbes 2021). Implementation environment, device, and measurement details can be found in (W. Unger, L. Babinkostova, M. Borowczak and R. Erbes 2021).

Deep Learning Power Analysis (DLPA). We applied a Deep Learning based Non-Profiled SCA proposed in (Timon 2019) by combining CPA-like hypotheses with Deep Learning training. The target function is $HW(Sbox(d_i \oplus k^*))$, where $(d_i)_{1 \leq i \leq N}$ are known random values and $k^* \in K$ is the fixed secret key value. We used two variants of DLPA, using MLP and CNN architectures as underlying neural networks. In addition, we used Hamming Weight (Hamming Weight is the number of non-zero symbols in a symbol sequence) labeling method to perform DLPA: $H_{i,k} = HW(V_{i,k})$, where $V_{i,k} = F(d_i, k)$ and $F(d_i, k^*) = Sbox(d_i \oplus k^*)$.

We conducted CPA, DLPA-MLP, and DLPA-CNN on 10 datasets of 345 power traces of GIFT-128, each with 1250 time samples. Our results, found in table **??**, show that both CPA and DLPA-CNN were able to successfully recover all round keys while DLPA-MLP only recovered the round-key in 6 out of 10 datasets.

Attack	Accuracy
CPA	100%
DLPA-CNN	100%
DLPA-MLP	60%

Table 1: Attacks performed on 10 different data sets, each with a different fixed key and 345 traces

During our experiments, the HW model for DLPA-CNN on GIFT-128 provided better results than in (Timon 2019) where the same model was weaker on Advanced Encryption Standard (AES) (Daemen and Rijmen 2003) (non-lightweight cryptographic standard). This illustrates the importance of the labeling method when conducting Non-Profiled SCA on different cryptographic schemes.

Implementation With De-synchronization. As the ChipWhisperer traces are by default almost perfectly synchronized, we added de-synchronization to study the efficiency of DLPA on GIFT-128 against de-synchronized traces. The implementation of de-synchronization involves offsetting the traces by adding random delays during encryption to simulate a clock jitter effect (E. Brier, C. Clavier, and F. Olivier 2004). The traces were de-synchronized through shifting each trace left or right by random values chosen in the interval [-25, 25]. The de-synchronization results for the three attacks are shown in Table **??**.

Attack	Accuracy
CPA	0%
DLPA-CNN	100%
DLPA-MLP	0%

Table 2: Attacks performed on 10 different de-synchronize datasets, each with a different fixed key and 345 traces

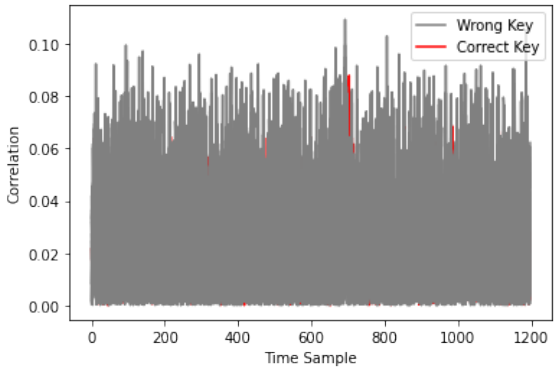

Figure 1: CPA conducted on de-synchronized traces. The correct key does not produce a singular spike in correlation or the highest correlation

As we can observe, DLPA-CNN attack, due to the translation-invariant property of its non-linear layers, is able to recover the round key with an accuracy of 100%, while CPA and MLP-DLPA are not. The inability of CPA to address de-synchronization can be seen in Table **??** and Figure 1 where the correlation for the correct key does not separate from the others. A similar issue happens with the DLPA-MLP that just overfits.

Acknowledgements

This research has been supported by the National Science Foundation under award #1950599.

References

Arribas, V. 2020. *Design and Verification of Side-Channel and Fault Attacks Countermeasures*. Ph.D. thesis, Katholieke Universiteit Leuven. Svetla Nikova and Vincent Rijmen (promotors).

Banik, S.; et al. 2017. GIFT: A Small Present – Towards Reaching the Limit of Lightweight Encryption. *IACR-CHES. ePrint Arch.*, 321–345.

Daemen, J.; and Rijmen, V. 2003. AES Proposal: Rijndael. *National Institute of Standards and Technology*.

E. Brier, C. Clavier, and F. Olivier. 2004. Correlation Power Analysis with a Leakage Model. *Cryptographic Hardware and Embedded Systems – CHES 2004*, 16–29.

E. Prouff, R. Strullu, R. Benadjila, E. Cagli, and C. Dumas. 2018. Study of Deep Learning Techniques for Side-Channel Analysis and Introduction to ASCAD Database. *Journal of Cryptographic Engineering*, 53.

P. Kocher, J. Jaffe, and B. Jun. 1999. Differential power analysis. *Advances in Cryptology – CRYPTO'99*, 388–397.

Timon, B. 2019. Non-Profiled Deep Learning-based Side-Channel attacks with Sensitivity Analysis. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2019(2): 107–131.

W. Unger, L. Babinkostova, M. Borowczak and R. Erbes. 2021. Side-channel Leakage Assessment Metrics: A Case Study of GIFT Block Ciphers. *J021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 236–241.