Master Internships on Deep Learning Side-Channel Security

Number of openings: 2/3

Start date: Flexible, anytime from March/April 2025.

Duration: 5-6 months.

Scientific context

After more than 20 years of research, Side-Channel Analysis (SCA) attacks are still one of the most critical vulnerabilities in embedded systems. By looking for correlations between processed data and physical, observable side effects of computing like power consumption, Electromagnetic (EM) emanations, or timing, SCA attacks have been traditionally directed to retrieve cryptographic keys from ciphers like AES. However, the increasing adoption of Machine and Deep Learning (ML, DL) is making Artificial Intelligence (AI) a new target. As these systems increasingly deal with sensitive data and control critical infrastructures and new vulnerabilities are reported, the **hardware/software security of ML/DL systems** is emerging as a key cybersecurity concern to build trustworthy AI-based systems [1, 2]. **SCA attacks to DNN implementations** enable the recovery of secret assets like models' structure, parameters, and private data inputs, which jeopardizes privacy and enables counterfeiting by reverse-engineering of models [3, 4] and the structure and dataflow scheduling of encrypted IP hardware accelerators [5]. Such side-channel-assisted information can also help adversaries fool systems more easily toward misclassifications. We are interested in both local SCA attacks to edge devices, highly exposed to attackers [6–9], and remote SCA attacks to cloud FPGAs [10, 11].

The traditional target of SCA has been a cryptographic key, so certain assumptions about the system runtime properties have usually been given for granted. One such assumption is that the system operates free of errors. However, to save energy, a new computing paradigm called **Approximate Computing (AxC)** aims at exploiting the tolerance to errors of certain applications by trading-off quality of results (e.g., precision or accuracy) with reduced usage of computational resources (energy, hardware, time), to allow building faster and less power-hungry computing systems. AxC techniques can be applied at different levels, from circuits all the way up to applications [12, 13]. Examples include (1) undervolting (reducing the power supply level even beyond the recommended margins of manufacturers), (2) approximate circuits, storage and memory, (3) software-level approximations like skipping computations through loop perforation.

Objectives of the internships

These internships are framed in the ANR JCJC project ATTILA[1] (2021–2025, young investigators grant from the French national research agency). The **objectives** are to **investigate the side-channel vulnerabilities of DL systems** and to **design secure implementations against SCA attacks**. The focus is either on SW implementations in microcontrollers or in HW accelerators in heterogeneous reconfigurable platforms (MPSoC-FPGAs). An initial step is the replication of existing attack/s from the literature, either to retrieve the model/architecture (hyperparameters), the parameters (weights, activation function), or the inputs. Although the focus is on physical side-channel vulnerabilities exploiting power consumption or EM emanations, the objectives can be adapted to explore other side-channel vulnerabilities, too. As the internship advances, different directions are possible and will be discussed with the students according also to their interests. The main ideas to explore revolve around:

[1] ATTILA: https://rsalvador.org/projects/ATTILA/

- **DNN implementations using AxC techniques.** Extend our current workflow and setup to implement DNN models in microcontrollers or FPGAs using AxC techniques and exploring frameworks like TinyML.
- **Evaluation of DNN side-channel security.** tudy the literature on standard side-channel evaluation methodologies and metrics (TVLA, SNR, etc...), and assess their adequacy in the context of DNN side-channel vulnerabilities.
- **Impact of DNN configurations and AxC techniques.** Study how different configurations, parameters and DNN implementations can affect the observable side channels. These can include:
  - Exact vs. AxC implementations at software or hardware level
  - Compiler optimizations
  - Microarchitectural features (cache configuration, multiple instruction issue, etc.)
- **Implementation and evaluation of countermeasures.** Study the existing countermeasures from the literature, implement and evaluate one of them, and/or study new approaches.

The position offers a clear path to **complete a PhD in an important emerging field** and the chance to **set up and develop their own research agenda** to postdoc candidates.

Candidates profile

We **welcome candidates with different backgrounds and interests**, e.g., on hardware and architecture (FPGAs, hardware security, secure accelerators and microarchitecture, microcontrollers) or on computer science/mathematics (side-channel analysis, cryptanalysis, artificial intelligence).

Master 1 or 2 students (or 4th/5th year Engineering) in Computer/Electrical Engineering, Embedded Systems, Electronics/Microelectronics or Computer Science. You should have a strong background in at least one of the following topics:

- Side-channel attacks, side-channel analysis and evaluation methodologies, cryptanalysis
- Other HW/SW security background
- Design for FPGAs and hands-on experience in prototyping and implementations
- HW or SW implementations of DNNs (FPGAs, microcontrollers, other accelerators/systems)
- ML/AI frameworks (TinyML, PyTorch, TensorFlow, TFLite...)

Other interesting skills to have:

- Programming in C/C++/Python
- Use of Linux/Git as development environment
- Good use of laboratory instruments (oscilloscopes, power supplies, etc.)

You can speak, write, and read English at a professional level (french language is not required).

Position details

Stipend: according to regulations, between 650-700 €/month

Supervisors

You will integrate the SUSHI team of IRISA/Inria in Rennes. We are part of a larger collaborative environment with researchers in Rennes and Lorient working on DL hardware/software security, so you will also work with members from the ASIC team of IETR and SHAKER of Lab-STICC.

Contacts of main supervisors:

- **Dr. Rubén Salvador**: `ruben.salvador@inria.fr`
- **Dr. Lorenzo Casalino**: `lorenzo.casalino@centralesupelec.fr`

How to apply

Please send us an email with the following information:

- Your CV
- Your Bachelor/Master transcripts (important to know your background)
- A motivational letter
- Any additional document/report or link to repositories that you can think can prove your experience

Application deadline: End of April 2025. Interviews will start as applications arrive so that the candidates might be selected before the deadline.

Please do not hesitate to contact us for further details and information.

References

[1] S. Mittal, H. Gupta, and S. Srivastava. "A Survey on Hardware Security of DNN Models and Accelerators". *J. Syst. Archit.* 117 2021, p. 102163. DOI: 10.1016/j.sysarc.2021.102163.

[2] V. Meyers, D. Gnad, and M. Tahoori. "Active and Passive Physical Attacks on Neural Network Accelerators". *IEEE Design & Test* 2023, pp. 1–1. DOI: 10.1109/MDAT.2023.3253603.

[3] M. Méndez Real and R. Salvador. "Physical Side-Channel Attacks on Embedded Neural Networks: A Survey". *Appl. Sci.* 11 15, 2021, p. 6790. DOI: 10.3390/app11156790.

[4] P. Horváth, D. Lauret, Z. Liu, and L. Batina. "SoK: Neural Network Extraction Through Physical Side Channels". *33rd USENIX Security Symposium (USENIX Security 24)*. 2024, pp. 3403–3422.

[5] C. Gongye, Y. Luo, X. Xu, and Y. Fei. "Side-Channel-Assisted Reverse-Engineering of Encrypted DNN Hardware Accelerator IP and Attack Surface Exploration". *2024 IEEE SP)*. IEEE Computer Society, Oct. 2023, pp. 1–1. DOI: 10.1109/SP54263.2024.00001.

[6] M. Isakov, V. Gadepally, K. M. Gettings, and M. A. Kinsy. "Survey of Attacks and Defenses on Edge-Deployed Neural Networks". *IEEE HPEC*. 2019, pp. 1–8. DOI: 10.1109/HPEC.2019.8916519.

[7] L. Batina, S. Bhasin, D. Jap, and S. Picek. "CSI NN: Reverse Engineering of Neural Network Architectures Through Electromagnetic Side Channel". *USENIX Security Symp.* 2019, pp. 515–532.

[8] R. Joud, P.-A. Moëllic, S. Pontié, and J.-B. Rigaud. "A Practical Introduction to Side-Channel Extraction of Deep Neural Network Parameters". *Smart Card Research and Advanced Applications*. Ed. by I. Buhan and T. Schneider. Cham: Springer International Publishing, 2023, pp. 45–65. DOI: 10.1007/978-3-031-25319-5_3.

[9] R. Joud, P.-A. Moëllic, S. Pontié, and J.-B. Rigaud. "Like an Open Book? Read Neural Network Architecture with Simple Power Analysis on 32-Bit Microcontrollers". *Smart Card Research and Advanced Applications*. Ed. by S. Bhasin and T. Roche. Cham: Springer Nature Switzerland, 2024, pp. 256–276. DOI: 10.1007/978-3-031-54409-5_13.

[10] Y. Zhang, R. Yasaei, H. Chen, Z. Li, and M. A. A. Faruque. "Stealing Neural Network Structure Through Remote FPGA Side-Channel Analysis". *IEEE Trans. Inf. Forensics Secur.* 16 2021, pp. 4377–4388. DOI: 10.1109/TIFS.2021.3106169.

[11] S. Moini, S. Tian, D. Holcomb, J. Szefer, and R. Tessier. "Power Side-Channel Attacks on BNN Accelerators in Remote FPGAs". *IEEE J. Emerg. Sel. Top. Circuits Syst.* 11.2 2021, pp. 357–370. DOI: 10.1109/JETCAS.2021.3074608.

[12] S. Mittal. "A Survey of Techniques for Approximate Computing". *ACM Computing Surveys* 48.4 Mar. 2016, 62:1–62:33. DOI: 10.1145/2893356.

[13] G. Armeniakos, G. Zervakis, D. Soudris, and J. Henkel. "Hardware Approximate Techniques for Deep Neural Network Accelerators: A Survey". *ACM Comput. Surv.* Mar. 2022. DOI: 10.1145/3527156.