Offer #2024-08418

Master Internships on Deep Learning Side-Channel Security

Contract type : Internship

Level of qualifications required : Master's or equivalent

Other valued qualifications : M1/M2 students (4thor /5th year Eng.) in Computer/Electrical Engineering, Computer Science, Embedded Systems, Electronics/Microelectronics

Fonction : Internship Research

Level of experience : Recently graduated

About the research centre or Inria department

The Inria center at the University of Rennes is one of eight Inria centers and has more than thirty research teams. The Inria center is a major and recognized player in the field of digital sciences. It is at the heart of a rich ecosystem of R&D and innovation, including highly innovative SMEs, large industrial groups, competitiveness clusters, research and higher education institutions, centers of excellence, and technological research institutes.

Context

The internships are expected to start around February/March and extend for up to 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, new vulnerabilities are reported, and the **hardware/software security of ML/DL systems** is emerging as a key cybersecurity concern for building 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, and (3) software-level approximations like skipping computations through loop perforation.

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 Symp. 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". IEEE S&P. 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. Springer, 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. Springer, 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.

Assignment

The **objectives** of these internships 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 HW accelerators in heterogeneous reconfigurable platforms (MPSoC-PGAs).

An initial step is **replicating existing attacks** 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 leakage sources, such as **microarchitectural side-channels**. As the internships advance, different directions are possible, and hence, specific activities will be discussed with the students according to their interests.

Main activities

Depending on the direction taken in each internship,**different lines of work** are possible:

- **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.** Study 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 the 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.

Skills

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 **technical skills** include:

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

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

Benefits package

- Subsidized meals
- Social, cultural and sports events and activities

General Information

- **Theme/Domain :** Security and Confidentiality
  Information system (BAP E)
- **Town/city :** Rennes
- **Inria Center :** Centre Inria de l'Université de Rennes
- **Starting date :** 2025-02-01
- **Duration of contract :** 6 months
- **Deadline to apply :** 2025-01-02

Contacts

- **Inria Team :** SUSHI
- **Recruiter :**
  Salvador Perea Ruben / ruben.salvador@inria.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply

Defence Security :
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST).Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :
As part of its diversity policy, all Inria positions are accessible to people with disabilities.