Towards Designing Machine Learning Attack Resistant PUFs

Elena Dubrova
KTH Royal Institute of Technology
SWEDEN

dubrova@kth.se

Keywords: Anti-tamper methods, PUF, modeling attack, side-channel attack, machine learning

EXTENDED ABSTRACT

In this talk, we will present several lightweight techniques for protecting Physical Unclonable Functions (PUFs) from Machine Learning (ML)-based modeling attacks. Due to the growing popularity of PUFs in applications such as secret key generation, challenge-response authentication, remote attestation, etc., many PUF-based constructions have been proposed. However, only a few of them are both ML modeling attack resistant and sufficiently lightweight enough to fit into low-end embedded devices. In the first part of the talk, we will present a lightweight PUF construction, CRC-PUF [1], in which input challenges are de-synchronized from output responses to make a PUF model difficult to learn. The de-synchronization is done by an input transformation based on a Cyclic Redundancy Check (CRC). By changing the CRC generator polynomial for each new response, we ensure that the success probability of recovering the transformed challenge is at most 2-86 for 128-bit challenges and responses.

In the second part of the talk, we will describe an alternative approach to combat ML-based modeling attacks, based on reconfigurability [2,3]. We will present a non-conventional arbiter PUF design that employs 4 × 4 switch blocks rather than 2 × 2 ones. A 4 × 4 switch block can be reconfigured in a variety of ways during the PUF's lifetime, allowing for frequent PUF updates. We will also demonstrate that a 4 × 4 arbiter PUF construction is significantly more area-efficient than 2 × 2 arbiter PUF constructions.

In the final part of the talk, we will discuss open issues concerning PUF security. We will discuss recently emerged profiled side-channel attacks that make use of deep learning [4,5]. In these attacks, a Neural Network (NN) is trained to learn features corresponding to \`0´s and \`1´s in the PUF's responses from power/electromagnetic traces captured from profiling devices. The resulting NN model is used to classify responses of the PUF implemented in the device under attack at run time.

REFERENCES

[1] Dubrova, E., Näslund, O, Degen, B., Gawell, A., Yu, Y. (2019), ”CRC-PUF: a machine learning attack resistant lightweight PUF construction”, in proceedings of IEEE European symposium on security and privacy workshops (EuroS&PW) of Workshop on Machine Learning for Cyber-Crime Investigation and Cybersecurity (MaL2CSec), Stockholm, Sweden, June 20, 2019, IEEE, pp. 264-271.

[2] Dubrova, E. (2018), “A reconfigurable arbiter PUF with 4 × 4 switch blocks”, in proceedings of IEEE 48th International Symposium on Multiple-Valued Logic (ISMVL), Linz, Austria, May 16-18, 2018, IEEE, pp. 31-37.

[3] Aknesil, C., Dubrova, E. (2021), “An FPGA implementation of 4 × 4 arbiter PUF”, in proceedings of IEEE 51st International Symposium on Multiple-Valued Logic (ISMVL), virtual event, May 25-27, 2018, IEEE, pp. 160-165.

[4] Yu, Y., Moraitis, M., Dubrova, E. (2020), "Why Deep Learning Makes it Difficult to Keep Secrets in the FPGA-as-a-Service Setting", in proceedings of Workshop in Dynamic and Novel Advances in Machine Learning and Intelligent Cyber Security (DYNAMICS), virtual event, Dec 7, 2020, ACM.

[5] Yu, Y., Moraitis, M., Dubrova, E. (2021), "Can Deep Learning Break a True Random Number Generator?", in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, issue 5, pp. 1710-1714.