Challenges in Deep Learning-Based Profiled Side-Channel Analysis

Stjepan Picek(✉)

Delft University of Technology, Delft, The Netherlands
stjepan@computer.org

Abstract. In recent years, profiled side-channel attacks based on machine learning proved to be very successful in breaking cryptographic implementations in various settings. Still, despite successful attacks even in the presence of countermeasures, there are many open questions. A large part of the research concentrates on improving the performance of attacks while little is done to understand them and even more importantly, use that knowledge in the design of more secure implementations. In this paper, we start by briefly recollecting on the state-of-the-art in machine learning-based side-channel analysis. Afterward, we discuss several challenges we believe will play an important role in future research.

1 Introduction

In side-channel analysis (SCA), the attacker exploits weaknesses in physical implementations of cryptographic algorithms [12]. This is possible by exploiting unintentional leakages in physical channels like power consumption [9] or electromagnetic radiation [19].

In profiled side-channel attacks, a powerful attacker has a device (the clone device) with knowledge about the secret key implemented and can obtain a set of profiling traces. From there, he builds a profiled model, which is then used to conduct an attack on another device (the device under attack). Consequently, profiled attacks have two phases (1) profiling phase where a model is constructed and (2) attack phase where the constructed model is used to attack the actual target device. Profiled SCA performs the worst-case security analysis as it considers the most powerful side-channel attacker with access to an open (since the keys are chosen/known by the attacker) clone device. The best-known profiled attack is the template attack, which is based on the Bayesian rule. Template attack is considered to be the most powerful attack from the information-theoretic point of view when the attacker has an unbounded number of measurements in the profiling phase [3]. To cope with certain statistical difficulties that can arise in template attack, there is a variant of it commonly known as the pooled template attack [4]. Finally, the third example of profiled attacks is the stochastic attack, which uses linear regression in the profiling phase [20].

These three techniques represent a standard set of techniques in profiled SCA. Besides these techniques, the SCA community also started using different

S. Bhasin et al. (Eds.): SPACE 2019, LNCS 11947, pp. 9–12, 2019.
https://doi.org/10.1007/978-3-030-35869-3_3

machine learning techniques. Common examples are the Naive Bayes [15], Support Vector Machines [7], Random Forest [10], and multilayer perceptron [6,13]. The multilayer perceptron algorithm (when having multiple hidden layers) also represents the first setting for deep learning-based attacks in profiled SCA. In 2016, Maghrebi et al. conducted a more detailed study of deep learning techniques in profiled SCA where they also used techniques like convolutional neural networks (CNN) or recurrent neural networks [11]. The reported results were in favor of CNNs, and from that time, a large part of the SCA community started to use CNNs, see, e.g., [2,17]. Such a direction seems to pay off as current state-of-the-art results suggest CNNs indeed perform very well and can break implementations protected with countermeasures [2,8,22].

2 State-of-the-Art and Future Challenges

We emphasize that we do not provide a complete overview of the state-of-the-art nor all related works tackling certain aspects of the future research directions we discuss. Rather, we concentrate on challenges we consider to be important and then offer more precise research questions within those.

Currently, the most explored research direction in machine learning-based SCA uses deep learning techniques like multilayer perceptron and convolutional neural networks to mount as powerful as possible attacks. A common setting is to use publicly available datasets (the more difficult dataset the more attractive target) and report the guessing entropy results (i.e., how many traces we require to break the target). There, we mention research by Kim et al. that showed how to add noise to the input to improve the performance of CNNs [8]. More recently, Zaid et al. proposed a methodology for CNN-based attacks where they achieved state-of-the-art results [22]. Some of their results are so good that it remains questionable whether truly better attacks on those datasets and in such scenarios are even possible (as minimal improvements in guessing entropy are not so relevant in practice). Still, there is room for improvements if we consider not only the number of measurements necessary to mount the attack but also to:

- Reduce the complexity of deep learning models. For example, Zaid et al. reported CNN models with much smaller number of parameters than commonly needed [22].
- Limit the number of measurements available to the attacker not only in the attack phase (which is usually done) but also in the training phase. By doing so, we force the attacker to use as powerful as possible deep learning models and at the same time, we reduce the computational complexity as the training phase would last shorter [16].
- Consider more difficult targets and more realistic settings. Indeed, a quite common procedure in profiled SCA research is to use only a single device for both profiling and attacking as well as to have the same key on both "devices". While this makes the setting easier for research, it also makes the results less reliable. Recent results indicate that settings using different devices and keys,

commonly known as portability settings, are significantly more difficult for machine learning attacks [1,5].

Next, despite strong results in deep learning-based SCA, we still do not understand much that is happening inside the deep learning process and as such, we do not know how to make the attacks even stronger. Common examples of questions one could ask are:

- How to know when to stop the training phase (as simply observing loss and accuracy is not necessarily revealing the SCA performance)?
- How to understand what did deep learning model learn and how different results one can expect from some other target?
- How to better connect the performance as measured by side-channel metrics and machine learning metrics?
- How to select the best deep learning architectures (from both performance and complexity perspectives) for certain scenarios and how to conduct good hyperparameter tuning?

We note there are several works partially considering such questions but the answers are far from complete [14,18,21].

Finally, while improving the performance of attacks is important, we must not forget that the end goal is to provide more security. As such, we should consider how to use the knowledge from the most powerful machine learning-based attacks to construct stronger countermeasures and how to use machine learning constructively in SCA (i.e., not only to attack).

References

1. Bhasin, S., Chattopadhyay, A., Heuser, A., Jap, D., Picek, S., Shrivastwa, R.R.: Mind the portability: a warriors guide through realistic profiled side-channel analysis. Cryptology ePrint Archive, Report 2019/661 (2019). https://eprint.iacr.org/2019/661
2. Cagli, E., Dumas, C., Prouff, E.: Convolutional neural networks with data augmentation against jitter-based countermeasures. In: Fischer, W., Homma, N. (eds.) CHES 2017. LNCS, vol. 10529, pp. 45–68. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66787-4_3
3. Chari, S., Rao, J.R., Rohatgi, P.: Template attacks. In: Kaliski, B.S., Koç, K., Paar, C. (eds.) CHES 2002. LNCS, vol. 2523, pp. 13–28. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-36400-5_3
4. Choudary, O., Kuhn, M.G.: Efficient template attacks. In: Francillon, A., Rohatgi, P. (eds.) CARDIS 2013. LNCS, vol. 8419, pp. 253–270. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08302-5_17
5. Das, D., Golder, A., Danial, J., Ghosh, S., Raychowdhury, A., Sen, S.: X-DeepSCA: cross-device deep learning side channel attack. In: Proceedings of the 56th Annual Design Automation Conference 2019, DAC 2019, pp. 134:1–134:6. ACM, New York (2019)
6. Gilmore, R., Hanley, N., O'Neill, M.: Neural network based attack on a masked implementation of AES. In: 2015 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), pp. 106–111, May 2015

7. Hospodar, G., De Mulder, E., Gierlichs, B.: Least squares support vector machines for side-channel analysis. Center for Advanced Security Research Darmstadt, pp. 99–104, January 2011
8. Kim, J., Picek, S., Heuser, A., Bhasin, S., Hanjalic, A.: Make some noise. Unleashing the power of convolutional neural networks for profiled side-channel analysis. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2019**(3), 148–179 (2019)
9. Kocher, P., Jaffe, J., Jun, B.: Differential power analysis. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 388–397. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48405-1_25
10. Lerman, L., Medeiros, S.F., Bontempi, G., Markowitch, O.: A machine learning approach against a masked AES. In: Francillon, A., Rohatgi, P. (eds.) CARDIS 2013. LNCS, vol. 8419, pp. 61–75. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08302-5_5
11. Maghrebi, H., Portigliatti, T., Prouff, E.: Breaking cryptographic implementations using deep learning techniques. In: Carlet, C., Hasan, M.A., Saraswat, V. (eds.) SPACE 2016. LNCS, vol. 10076, pp. 3–26. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49445-6_1
12. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks: Revealing the Secrets of Smart Cards. Springer, New York (2006). ISBN 0-387-30857-1. http://www.dpabook.org/
13. Martinasek, Z., Zeman, V.: Innovative method of the power analysis. Radioengineering **22**(2), 586–594 (2013)
14. Masure, L., Dumas, C., Prouff, E.: A comprehensive study of deep learning for side-channel analysis. Cryptology ePrint Archive, Report 2019/439 (2019). https://eprint.iacr.org/2019/439
15. Picek, S., Heuser, A., Guilley, S.: Template attack versus Bayes classifier. J. Cryptogr. Eng. **7**(4), 343–351 (2017)
16. Picek, S., Heuser, A., Guilley, S.: Profiling side-channel analysis in the restricted attacker framework. Cryptology ePrint Archive, Report 2019/168 (2019). https://eprint.iacr.org/2019/168
17. Picek, S., Heuser, A., Jovic, A., Bhasin, S., Regazzoni, F.: The curse of class imbalance and conflicting metrics with machine learning for side-channel evaluations. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2019**(1), 209–237 (2019)
18. Picek, S., Samiotis, I.P., Kim, J., Heuser, A., Bhasin, S., Legay, A.: On the performance of convolutional neural networks for side-channel analysis. In: Chattopadhyay, A., Rebeiro, C., Yarom, Y. (eds.) SPACE 2018. LNCS, vol. 11348, pp. 157–176. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-05072-6_10
19. Quisquater, J.-J., Samyde, D.: ElectroMagnetic analysis (EMA): measures and counter-measures for smart cards. In: Attali, I., Jensen, T. (eds.) E-smart 2001. LNCS, vol. 2140, pp. 200–210. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45418-7_17
20. Schindler, W., Lemke, K., Paar, C.: A stochastic model for differential side channel cryptanalysis. In: Rao, J.R., Sunar, B. (eds.) CHES 2005. LNCS, vol. 3659, pp. 30–46. Springer, Heidelberg (2005). https://doi.org/10.1007/11545262_3
21. van der Valk, D., Picek, S.: Bias-variance decomposition in machine learning-based side-channel analysis. Cryptology ePrint Archive, Report 2019/570 (2019). https://eprint.iacr.org/2019/570
22. Zaid, G., Bossuet, L., Habrard, A., Venelli, A.: Methodology for efficient CNN architectures in profiling attacks. Cryptology ePrint Archive, Report 2019/803 (2019). https://eprint.iacr.org/2019/803