SIDE CHANNEL ATTACKS TO RECOVER POWER TRACE USING DEEP LEARNING TECHNIQUE

Harisu Abdullahi Shehu[1], Md. Haidar Sharif[2], Rabie A. Ramadan[2], Adwan Alownie Alanazi [2]
[1]Pamukkale University, Turkey
[2]University of Hail, Kingdom of Saudi Arabia
Email: harisushehu@gmail.com, md.sharif@uoh.edu.sa, ra.ramadan@uoh.edu.sa, a.alanazi@uoh.edu.sa
(Presented at CSC, 2019, SA.)

ABSTRACT: *Side channel attack is an area of research that has received much attention in recent years. It imposes serious threats to the security of a system due to the information leaked from it. While deep learning based approaches have widely been used and have outperformed other existing machine learning approaches in the image and automatic speech recognition, only a few numbers of researches have been carried out to recover secret keys using deep learning. In this paper, we propose a deep learning approach to recover secret keys from power traces. We can recover up to 72.5% of the power traces using our method.*

Keywords: Deep learning, Machine learning, Power trace, Side channel attacks.

1. INTRODUCTION

Information leakage from electronic devices is now well known and has become a main topic of concern. The leaked information includes, but not limited to power, electromagnetic and sound. Due to the fact that this leakage depends on internally secret keys, an efficient key-recovery attack to reveal sensitive data [1-3] may be formed by an attacker to endanger the security of a system. Since the first public awareness of these threats [4], a lot of effort has been made to make research on side channel attacks and to develop corresponding countermeasures. Small hardware implementations are vulnerable to a range of side-channel attacks [5]. Electromagnetic radiation, power consumption, and timing are the three commonly vectors that leak information about computation and data on a chip [6]. Knowing that the architecture of a device is vulnerable to side-channel exploitation helps in deciding whether or not to store data on devices with similar processor characteristics and memory or to execute unprotected sensitive computations [7]. Side channel attack may be distinguished into two different classes [8] namely: (i) The profiling side channel attack which are the most powerful type of side-channel attacks [9]; (ii) Non-profiling side channel attack that corresponds to the much weaker side channel attack in which the attacker has only access to the physical leakage captured on the target device. Most cryptographic algorithms are implemented as cryptographic coprocessors [10, 11] or an application specific integrated circuits [12, 13] to meet the main requirements of demanding high throughput [14].

In this paper, we propose a deep learning (DL) based approach to solve the problem of side channel attack to recover keys from power traces on the standard evaluation board SASEBO-GII using an unmasked dataset.

The rest of the paper is organized as Section 2 illustrates related works; Section 3 explains our approach; Section 4 reports experimental results and discussion followed by few clues for further study, and Section 5 concludes the paper.

2. RELATED WORKS

The alternative attack that leads to the discovery of cryptographic keys technique in the physical implementation of a cryptographic device is called Side Channel Analysis (SCA). Hospodar et al. [15] used machine learning (ML). They have used a kernel-based learning algorithm known as the least-squares support vector machine and provided a number of power trace to the classifier to teach it. One unseen power trace is presented to the classifier at a time of the testing phase, and the result shows that the choice of parameters of the machine learning technique strongly impacts the performance of the classification.

On the other hand, the time instants and the number of power traces do not influence the result in the same proportion. This study demonstrates two fundamental techniques of power analysis attack which are the differential power analysis (DPA) which usually requires a high dimension power trace [16] and the correlation power analysis or the chosen-plaintext attack (CPA) against the Arduino Uno microcontroller. The differential of means attack is the kind of the DPA implemented and the CPA is implemented by building a power model of the device using the Hamming Weight Power Model method. While the experimental results find that CPA produces a result which is easier to interpret from an analytical perspective, it was also found out that both DPA and CPA are viable against the Arduino Uno [17].

Although dot matrix printers are outdated, it continues to play an important role in businesses where confidential information is processed. Backes *et al.* [18] used ML techniques for acoustic side-channel attacks on dot matrix printers. They presented a novel attack that recovers what a dot matrix printer processing English text is printing based on the recorded sound with the microphone placed at a distance that is very close to the printer. They have used a combination of ML, audio processing, and speech recognition techniques including spectrum features, Hidden Markov Model, and linear classification. With the microphone placed at a distance that is 10 cm from the printer, experimental results showed that the attack recovers up to 72% of printed words and up to 95% if the contextual knowledge about the text is assumed.

Recent research has investigated new profiling approaches mainly by applying ML technique. Maghrebi *et al* [1] used DL inside channel context to target both protected and unprotected cryptographic implementations. They tried to build an accurate profiling using DL which leads to an efficient and successful side channel key recovery attack. Their results have shown that the DL based attacks are more efficient than the ML-based and template attacks when it comes to targeting either unprotected or masked cryptographic implementations.

3. PROPOSED METHODOLOGY

This section describes our approach. We utilize the Convolution Neutral Network (CNN) to recover the power traces. As can be seen in Fig. 1, in the first phase, the data are collected and some of them are fed to the CNN training module. Afterward, the testing data are fed to CNN to check on the results.

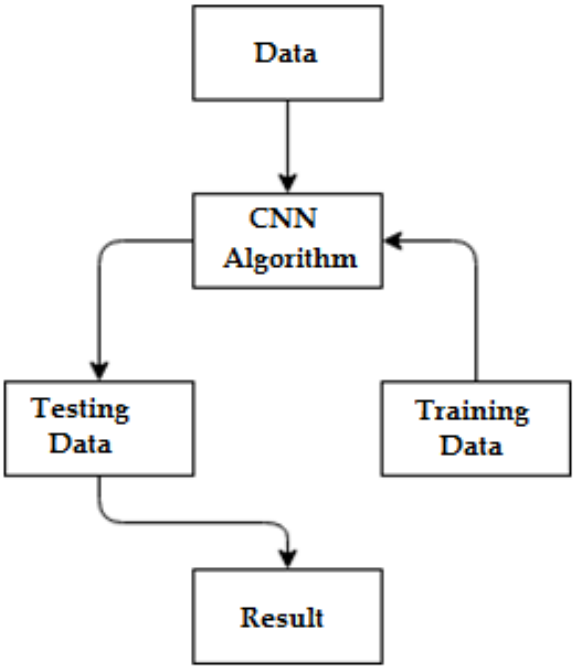


Figure 1. Flow Chart of the CNN Algorithm

3.1 Data Collection

Deep learning models need to be trained with a lot of data to perform well. Consequently, a very large amount of power traces data needs to be collected and to be used in both training and testing the model. But since producing the real-world data is obviously time-consuming, costly, and potentially inconvenient to collect, researchers have now started to release their datasets publicly to enable other researchers to compare their approach against a common benchmark dataset.

In this paper, we have used a benchmark dataset consisting of up to 100000 unmasked power traces [19]. A link to download the dataset has been sent via email after the request approval. The dataset is divided into two parts: (i) Each part consists of 50000 data making it a total of 100000 for the two unmasked datasets; (ii) Each part is of equal size in terms of memory with each consisting of up to 62.8 MB making it a total of 125.6 MB for the whole 100000 unmasked data.

3.2 The Usage of Convolution Neutral Network (CNN) Algorithm

Upon gathering a very large amount of dataset, the next step is to train the DL model. In general, training a DL model can take a very long time ranging from days to weeks. This process can significantly be speed up and the training time can be cut down from days to hours if the Graphical Processing Unit (GPU) is being used [20]. However, in this research, the GPU was not used to reduce the amount of training time since the data used in the study were not much in terms of memory.

In this paper, we have used CNN algorithm and fed the algorithm with a high amount of data to obtain a better accuracy [21] result. We have trained a sum of 95000 data records using an 8GB RAM HP 64-bit computer with an Intel Core i5-7200U CPU utilizing Windows 10 Pro. It took an approximate of 3 hours to finish the training process.

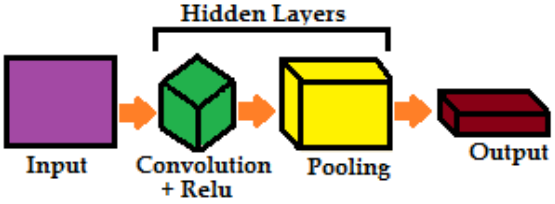


Figure 2. Illustration of the CNN layers.

Fig. 2 shows an illustration of the convolution neural network layers used in this research. There exist two hidden layers which are the convolution + relu and the pooling layers. The algorithm is fed with the unmasked dataset as input. Several computations are performed in the hidden layers to recover the traces.

3.3 Result Analysis

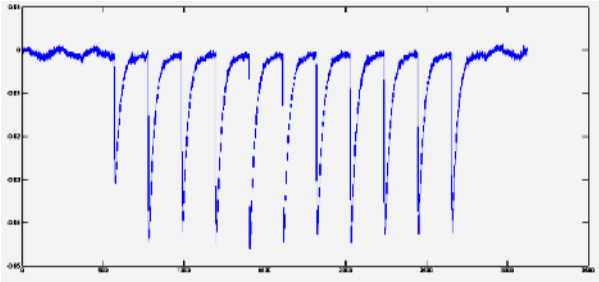
After successfully training the DL model, some set of data from the collected data were given to the developed model as the testing dataset. Accuracy of the developed model was determined according to the given testing data.

4. EXPERIMENTAL RESULTS

We have downloaded a benchmark dataset consisting of 100000 power traces from [19] and used it together with the SASEBO-GII standard evaluation board to recover power traces. One example of the unmasked power traces downloaded is presented in Fig 3 where the loading of plaintext into register represents the first peak and the following ten peaks represent the rounds in advanced encryption standard. We have divided the downloaded data into two different classes - one for training whereas the other class was used for testing the algorithms using the FPGA Sasebo-GII evaluation board. We have used only 5000 out of 100000 available data for testing the algorithm. The 100000 data used in this research is only 125.6 MB and 5000 of the data was used for testing whereas the remaining 95000 were used for training the CNN algorithm.

We were able to recover up to 3625 out of the 5000 data. Henceforth the achievement approaches to an accuracy of (3625/5000 =) 72.5%. We have considered this result is satisfactory since CNN is new to this field. In addition, CNN requires large training sets to produce better results. Here, we have set a sub-standard for other researchers working on side channel attacks using AI approaches. Hence, we believe that it is possible to obtain a better result with further improvement by updating the CNN algorithm in our future work. The researchers of [19] that performed the same research on the SASEBO-GII evaluation board using a different methodology were able to recover 100% accuracy. However, regular methods were used for that accomplishment. From our point of view, we believe, it is the first time DL is used in this field and further improvement is possible.

More data would be used to train the DL model as DL requires to be trained with enough quality data in order to perform well. In addition, the CNN algorithm will be updated by possibly adding more hidden layers to it so as to obtain a better accuracy result than the one already achieved. A complete statistical analysis [22] along with computational complexity [23] would be studied in the long



run to compare the performance of miscellaneous state-of-the-art approaches.

Figure 3. Example of unmasked power trace on FPGA, Sasebo-GII [19].

6. CONCLUSION

We aimed to solve a side channeling attack problem by recovering secret keys from power traces using deep learning and convolution neural network. We obtained up to 72.5% of the power traces used on Sasebo-GII are being recovered. Future work would include training the DL model with more data for performing well.

8. REFERENCES

[1]. Maghrebi H., Portigliatti, T., Prouff E., "Breaking Cryptographic Implementations using Deep Learning Techniques", International Conference on Security, Privacy, and Applied Cryptography Engineering, pp. 3-26, (2016).

[2]. Zhou Y., Feng D., "Side-Channel attacks: ten years after its publication and the impacts on cryptographic module security testing", *IACR Cryptology ePrint Archive*, 388, (*2005*).

[3]. Standaert, F.-X, "Introduction to Side-Channel Attacks", *UCL Crypto Group, Place du Levant 3, B-1348 Louvain-la-Neuve,* (2010).

[4]. Kocher, P.C., Jaffe, J., Jun, B., "Differential power analysis", *In: Crypto 99—Advances in Cryptology. LCNS, vol. 1666, pp. 388– 397,* (1999).

[5]. Oswald et al., Side-Channel Analysis Resistant Description of the AES S-Box. (2005).

[6]. Lerman L. & Bontempi G., Markowitch O., "Side-channel attack an approach based on machine learning", *Université Libre de Bruxelles*, (2011).

[7]. Banerjee U, Ho L, Koppula S., "Power-Based Side-Channel Attack for AES Key Extraction on the ATMega328 Microcontroller", *Computer Systems Security*, (2015).

[8]. Picek, S., Heuser, A., Jovic, A., Ludwig, S. A., Guilley, S., Jakobovic, D., & Mentens, N., "Side-channel analysis and machine learning: A practical perspective", *International Joint Conference on Neural Networks (IJCNN), 4095 – 4102,* (2017).

[9]. Standaert, F. X., Koeune, F., Schindler, W., "How to Compare Profiled Side-Channel Attacks?" *In: Abdalla, M., Pointcheval, D., Fouque, P.A., Vergnaud, D. (eds.) ACNS. LNCS, 485–498,* (2009).

[10]. J. Goodman and A.P. Chandrakasan, "An Energy-Efficient Reconfigurable Public-Key Cryptography Processor", *IEEE Journal of Solid-state Circuits*, 1808- 1820, (2001).

[11]. Correnf Datasbeet-Product Brief of Corrent Packet Armor CR7110 Security Processor, (2003). http://www.corrent.com/pdfs/CR7110~%20Pbrief~v3.p

[12]. H. Technology, Datasheet High Performance DES and Triple DES core for ASIC, (2003).

[13]. I. Verbauwhede, P. Schaumont and K. Kuo, "Design and performance testing of a 2.29 Gb/s Rijndael processor", *IEEE Joumal of Solid-State. Circuits*, pp. 569- 572, (2003).

[14]. B. Yang, K. Wu, R. Karri, "Scan Based Side Channel Attack on Dedicated Hardware Implementations of Data Encryption Standard" *International Conferce on Test*, pp. 339-344, (2004).

[15]. Hospodar, G., Gierlichs, B., De Mulder, E., Verbauwhede, I., and Vandewalle, J., "Machine learning in side-channel analysis: a first study", *Journal of Cryptographic Engineering*, vol. 1(4), pp. 293–302, (2011).

[16]. Hogenboom J., "Principal Component Analysis and Side-Channel Attacks", *Master Thesis, Radboud University Nijmegen, Thesis No. 634*, (2010).

[17]. Lo O., Buchanan W. J., Carson D., "Power analysis attacks on the AES-128 S-box using differential power analysis (DPA) and correlation power analysis (CPA)", *Journal of Cyber Security Technology*, vol. 1, pp. 88-107, (2017).

[18]. Backes, M., Dürmuth, M., Gerling, S., Pinkal, M., Sporleder, C., "Acoustic side-channel attacks on printers" *In: USENIX, p. 20 USENIX Association, USA*, (2010).

[19]. Tescase: http://tescase.coe.neu.edu/?current_page=POWER_TRACE_LINK

[20]. Mathworks:https://www.mathworks.com/discovery/deep-learning.html

[21]. M. H. Sharif, "An eigenvalue approach to detect flows and events in crowd videos," *Journal of Circuits, Systems and Computers*, vol. 26, no. 07, p. 1750110, (2017).

[22]. H. Kusetogullari, M. H. Sharif, M. S. Leeson, and T. Celik, "A reduced uncertainty-based hybrid evolutionary algorithm for solving dynamic shortest-path routing problem," *Journal of Circuits, Systems, and Computers*, vol. 24, no. 5, 2015.

[23]. M. H. Sharif, "A numerical approach for tracking unknown number of individual targets in videos," Digital Signal Processing, vol. 57, pp. 106 – 127, 2016.