A Survey of Prototype Side-channel Attacks Based on Machine Learning Algorithms for Cryptographic Chips

Dongxin Guo¹*, Kaiyan Chen¹, Xiaoyang Hu¹, Yanhai Wei¹, Jianlong Li¹

¹Department of information Engineering Mechanical Engineering College
Shijiazhuang, China

*Corresponding author e-mail: 810644463@qq.com

Abstract. Cryptographic chips have been widely used in the field of digital security. With the continuous development of side-channel attacks, the security of cryptographic chips has attracted more and more attention. This paper mainly reviews the most distinguishing template attacks in current side-channel attacks, introduces the basic steps of template attack implementation, analyzes and discusses the existing template attacks, and finally focuses on the research progress of template attacks, especially the current the research status of template attacks based on machine learning and deep learning algorithms lays the foundation for further research.

1. Introduction

The current cryptographic chip is very versatile and plays a very important role in protecting the security of critical information. Its role in the military field is particularly irreplaceable, and it protects important military information and intelligence. Traditional encryption algorithms use mathematically difficult problems to ensure the security of the Cryptographic chip.

However, with the proposed side-channel attack, especially the template attack distinguisher, the information security of the cryptographic chip poses a huge challenge. side-channel analysis is performed by exploiting the power consumption, electromagnetic, time and other side-channel information leaked by the cryptographic chip during the encryption process. The template attack is the attack capability of the side-channel distinguisher (eg, simple attack, differential attack, correlation attack). Relatively strongest method, so that the encryption key of the device is cracked more efficiently, and the relevant encrypted information is obtained.

Therefore, the main content of this paper includes the following three parts. The first part describes the implementation of the template attack side-channel distinguisher in the current side-channel attack. The second part discusses the problems and solutions of the traditional template attack. Finally, Traditional template attacks and the progress of template attacks based on machine learning and deep learning are summarized to lay the foundation for further research.

2. Template Attack Implementation

Since the introduction of the template attack [1] in 2002, it has received great attention from researchers. The template attack is one of the side-channel distinguishers. Since the template attack can obtain more information from the side-channel information, it is considered to be the strongest side-channel distinguisher. The traditional template attack characterizes the side-channel signal trajectory features by using a multivariate Gaussian distribution. The template attack can be divided
into three phases, the first phase is the extraction and selection of feature points; the second phase is template construction; the third The stage is template matching. The following is an overview of the various stages of the template attack.

2.1 Feature point selection and extraction

Before the template attack, feature extraction of the collected energy trajectory is required, and the maximum correlation point of the feature is selected. It is beneficial to improve the accuracy and attack efficiency of template construction. Here are five main feature selection methods:

- **SOD (Sum Of pairwise Differences) method.** Chari [1] proposed the SOD method in 2003, which is a relatively simple feature point selection method. The calculation of the SOD method is shown in Equation 1:

  \[ S = \sum_{j>i=1}^{N} (t_i - t_j) \]  

- **SOS (Sum Of Squared pairwise Differences) method.** Christian et al. proposed an SOSD improvement by analyzing the SOD method. In the method, because the collected data has positive and negative attributes, the SOD method cannot be processed well in the calculation process. The SOSD method adds a square operation based on the SOD method. The calculation method is as shown in Equation 2:

  \[ S = \sum_{j>i=1}^{N} (t_i - t_j)^2 \]  

Similar to the SOSD method, the time at which a larger value is selected is taken as the position at which the feature point is selected.

- **SOST (Sum Of Squared pairwise T-differences) method.** Gierlichs [2] et al. proposed a feature point selection method SOST based on T-test to optimize the SOD and SOSD methods. The calculation method of the SOST method is shown in Equations 3 and 4:

  \[ T = \frac{m_i - m_j}{\sqrt{\frac{\sigma_i^2}{N_i} + \frac{\sigma_j^2}{N_j}}} \]

  \[ S = \sum_{j>i=1}^{n} \left( \frac{m_i - m_j}{\sqrt{\frac{\sigma_i^2}{N_i} + \frac{\sigma_j^2}{N_j}}} \right)^2 \]

Equation 3 is the calculation method of T-test, formula 4 is the SOST feature point extraction calculation mode, \( m \) is the curve set mean vector based on a certain intermediate value category, indicating the corresponding variance curve, and \( N \) is the corresponding energy curve number. Similar to the above method, the time at which a larger value is selected is taken as the position at which the feature point is selected.

- **Pearson correlation coefficient method.** Before you build a template, you need to find the correlation with the median in the side-channel track. The position of the largest point, thereby determining the selection range of the feature points, and efficiently establishing an accurate template. The main role of the Pearson correlation coefficient is to calculate the linear correlation between the two data sets. See Equation 5 for the calculation method:

  \[ \rho(i) = \frac{Cov(t_i, v)}{\sqrt{Var(t_i) \cdot Var(v)}} \]
Where $t$ denotes a vector composed of points corresponding to the point $i$ of the energy trace at a certain moment, and $v$ denotes a corresponding intermediate value. The time at which the correlation coefficient is larger is selected as the position at which the feature point is selected.

- Principal Component Analysis (PCA) method. Principal component analysis is a linear mapping method, the principle is to reduce the data concentration. Sign similar data to reduce the number of feature points. In the specific application process, it can be divided into four steps; the first step is to preprocess the collected raw data; the second step is to obtain the eigenvalues and feature vectors for the preprocessed data, and sort them; The variance contribution rate is selected, and the feature vector is selected. The fourth step is to calculate the feature principal component value of the data set, and obtain the feature point set used for template construction and testing.

2.2 Template construction

In the template construction process, the mean vector and the covariance matrix of the multivariate positive distribution corresponding to each template are estimated. The size of the covariance matrix is proportional to the square of the number of points on the energy trace. Obviously, it is necessary to find a feature point. Strategy, the number of feature points is NIP. A feature point is a point that contains the most information about the instruction being drawn. The following describes three common template construction [3] strategies:

- Data and key pair template construction. The first strategy. That is, to build a template for each data and key pair. Therefore, the feature points used to construct the template are all points in the energy trace that are related to the key pair.

- Intermediate value template construction. The second strategy. The template is built by some suitable function. Therefore, the feature points used to construct the template in the energy trace are the energy traces corresponding to all operations involving the function.

- Template construction based on energy model. The commonly used energy models for side channel attacks are the Hamming weight model and the Hamming distance model. Essentially, it belongs to the Hamming distance model, which assumes that the device has the same energy consumption in $1 \rightarrow 0$ and $0 \rightarrow 1$ data conversion, and has the same energy consumption in $0 \rightarrow 0$ and $1 \rightarrow 1$ data conversion. influences.

2.3 Template matching

In the actual template attack, there are many difficulties in the template matching phase. These difficulties are related to the covariance matrix. First, the size of the covariance matrix depends on the number of feature points. Obviously, the number of feature points must be carefully chosen; secondly, the covariance matrix may be "morbid". This means that the covariance matrix will encounter numerical problems when inverting, and the inversion is necessary to calculate Equation 6. In addition, the indices in exponential operations tend to be small, which often leads to more numerical problems.

$$p(T^i;(m,C)_{\text{hmr}}) = \frac{\exp\left(-\frac{1}{2}\cdot(T^i - m)^\top \cdot C^{-1} \cdot (T^i - m)\right)}{\sqrt{(2 \cdot \pi)^{N_v} \cdot \text{det}(C)}}$$

(6)

In principle, template matching is the magnitude of the probability that a given energy trace is evaluated using Equation 6 for a given template. The template that produces the highest probability will reveal the correct key.

In order to avoid the exponential operation, the logarithm of the formula 6 is usually obtained to obtain the formula 7. Thus, a template that minimizes the absolute value of the probability log can reveal the correct key:

$$\ln p(T^i;(m,C)) = -\frac{1}{2} \cdot \left(\ln(2 \cdot \pi)^{N_v} \cdot \text{det}(C) + (T^i - m)^\top \cdot C^{-1} \cdot (T^i - m)\right)$$

(7)

In order to avoid numerical problems when inverting the covariance matrix, replace the covariance
matrix with the unit matrix. Essentially, this means that there is no need to consider the covariance between points. A template consisting only of mean vectors is called a simplified template. Replacing the covariance matrix with the unit matrix simplifies the multivariate normal distribution. See Equation 8 for details:

\[
p(T^i; m) = \exp\left(-\frac{1}{2}(T^i - m)^\top(T^i - m)\right) / \sqrt{(2\pi)^{N_p}}
\]

In order to avoid the problems encountered in the exponential operation, the logarithm of the formula 8 is also taken. The simplified calculation of the probability log is as shown in Equation 9:

\[
\ln p(T^i; m) = -\frac{1}{2}(\ln(2\pi)^{N_p} + (T^i - m)^\top(T^i - m))
\]

3. Problems with Template Attack and The Solution Direction

Since the introduction of the template attack, it has received extensive attention from researchers because of its strong distinguishing ability. What we call the traditional template attack is based on the assumption that the captured energy trace conforms to the multivariate Gaussian distribution, and the side-channel distinguisher (such as simple energy analysis (SPA), correlation energy analysis (CPA) can be obtained from the side-channel information), Differential Energy Analysis (DPA), etc.) are more information and are therefore considered to be the strongest side-channel distinguishers.

However, in the specific implementation stage, due to problems such as singular matrix in the calculation process, if the side-channel energy trajectory acquisition is limited, it may result in the inability to use more feature points to construct the template. At the same time, the implementation of the traditional template attack is based on the Gaussian multivariate distribution, which is a probability density function. In most cases, it is more consistent with the leaked data. But there is no rigorous proof at the moment, so is there a better way to characterize the captured data and improve the accuracy of the characterization.

In view of the above problems, there are mainly two ways to solve the problem. One way is to modify or remove the part that is easy to cause mathematical calculation problems by simplifying the calculation formula, but this situation usually leads to a decrease in the effect of the template attack. Another way is to use template learning based on machine learning and deep learning to avoid template deficiencies when dealing with high-dimensional feature data. On the other hand, it has strong feature capture capability and thus improves the template. The efficiency of the attack. The following is a review of the research progress of traditional template attacks and template attacks based on machine learning and deep learning in side-channel attacks.

4. Template Attack Classification

4.1 Traditional template attack

Traditional template attacks were first proposed by Chari [1] and others, and take advantage of the fact that the amount of energy consumed is related to the data being processed. And the method uses a multivariate normal distribution to characterize the acquired data features.

Based on the traditional template attack field, the researchers have done a lot of work. In the aspect of feature point, the SOD feature extraction method proposed by Chari et al. [1] is the first one. Christian Rechberger et al. optimized the SOD method and proposed the SOSD feature extraction method [4], which avoids the influence of data symbols on the calculation. Subsequently, F.-X. Standaert proposed the feature extraction method of principal component analysis and Fisher linear discriminant [5]. The main principle is to reduce the number of feature points and improve the attack efficiency by removing the feature points with higher similarity. Gierlichs [2] and others proposed the SOST method, which is a feature extraction method based on the T test, which can also better extract the feature data in the side-channel signal trace.
Subsequently, D Agrawal and JR Rao et al. [6] proposed an enhanced DPA template attack, which combines traditional DPA with single-bit template attack to successfully implement key attack. Rechberger [7] and others continued Chari's work, using a template attack to attack the RC4 encryption algorithm running on the microcontroller. Archambeau et al. [8] studied the feature point selection of traditional template attacks, selected the minimum spacing and the number of feature points, and proposed a method based on the execution of template attacks in the main subspace of the trajectory, and used RC4 encryption algorithm. Verification, experiments show that the attack greatly reduces the amount of encrypted information needed. Gierlichs et al. compared the effects of classical template attacks and Stochastic Methods attacks on software-based AES algorithms. Oswald and paar [9] used a template attack to crack the 112-bit key of the 3DES encryption algorithm executed on the contactless smart card. The experimental results show that the template attack poses a serious threat to the security of the DESFire MF3ICD40 smart card. Choudary and Kuhn [10] experiments have shown that changes caused by the use of different devices or different acquisition work on the same device have a great impact on the template attack. For this phenomenon, the author uses Fisher linear discriminant analysis and principal component analysis to The performance of the template attack is optimized. Fan [11] added a normalization operation to the traditional template attack, proposed a normalized template attack, and compared it with the traditional template attack. The experimental results show that the normalized template attack is more effective.

Researchers have done a lot of research in the field of traditional template attacks, and have achieved some good results, but still need further optimization and improvement in terms of specific calculation and efficiency.

4.2 Template attack based on machine learning and deep learning algorithms

The current template attacks based on machine learning and deep learning algorithms mainly include: using neural networks, random forests, support vector machines, deep neural networks and other pattern classifiers. In fact, the machine learning method does not make assumptions about the probability density distribution of the data. For example, the random forest establishes a decision tree set, and the data set is classified by the voting system [12]. The support vector machine based attack uses hyperplane clustering to distinguish the data. Sets, neural network-based attacks are mainly based on constantly adjusting the hyperparameters in the neural network to fit the data set that needs to be classified. The main disadvantage of the traditional template attack is that it can't effectively deal with high-dimensional data, and the machine learning based template attack can solve this problem well.

The simplest neural network model is the perceptron, and the current multi-layer perceptron has been successfully applied in the field of side-channel attack key recovery. Gilmore et al. [13] used a neural network to crack the key of the masked AES energy trajectory provided by the DPA-V4 competition. The attack idea is to first use the neural network to crack the mask, and then use the neural network to crack the algorithm key. The attack can be implemented based on the intermediate value obtained by the enemy after the mask operation, but in reality, the assumption is Not always achievable.

A random forest is made up of many decision trees, and a decision tree is a tool that uses binary rules to classify data. In the classification process, each tree is trained using different categories of data in the training data set, and the output is determined based on the results of all decision tree voting. Random forests have been successfully applied in the field of side-channel attacks [12]. Lerman et al. used principal component analysis to extract the feature data needed to construct the template, used the random forest method to construct the template, and cracked the DES algorithm. The results showed that the key search range was reduced by nearly half compared with the traditional template attack.

The support vector machine [14] is a linear classifier that not only finds a hyperplane to separate data classes, but also optimizes the problem of maximizing different class boundaries. To handle non-linearly separable data sets, you can use a kernel function that maps data to a higher-dimensional feature space. In the field of side-channel attacks, researchers have used support vector machines to perform key cracking on encryption algorithms without protection strategies and encryption algorithms with protection strategies. Heuser and Zohner [15] have proved that when collected data sets When the noise is low, the template attack based on support vector machine is better than the
traditional template attack. The author Gabriel uses the support vector machine for template attack and compares it with the traditional template attack. The comparison result is that the Hamming weight of the attack object has a greater impact on the attack effect.

In addition, Lerman [16] et al. used machine learning algorithms for template attacks, and compared classic template attacks with template learning based on machine learning algorithms. The conclusion is that in the case of black boxes, machine learning template attacks are used instead. Classic template attacks are possible. AES algorithm for masks. They also studied that the machine learning based template attack has a significant improvement in both time and space compared to the traditional template attack.

According to the data that can be collected at present, the method of side-channel template attack based on deep neural network has been studied in a few cases in recent years. Maghrebi [17] of Safran Group used classic template attack method and deep learning respectively. The technical method performs key cracking on the DPA Contest game data. The experimental results show that the template attack effect based on deep neural network is more efficient than the classic template attack.

References
[1] Chari, S., Rao, J.R., Rohatgi, P.: “Template attacks”. In: Kaliski Jr., B.S., Ko, c, C ,K., Paar, C. (eds.) CHES 2002. LNCS, vol. 2523, pp. 172–186. Springer, Heidelberg. 2003.
[2] Gierlichs B, Lemke-Rust K, Paar C. Templates vs. stochastic methods[C]// International Conference on Cryptographic Hardware and Embedded Systems. Springer-Verlag, 2006:15-29.
[3] Mangard S,Oswald E,Popp T. energy analysis attack[M]. Dengguo Feng,Yongbin Zhou,Jiye Liu,etc. beijing: 科学出版社, 2010.
[4] C.Rechberger and E.Oswald.“Practical Template Attacks”.In Workshop on Information Security Applications — WISA 2004,August 2004.
[5] C. Arachambeau, E. Peeters, F. -X. Standaert, J. -J. Qisquater, “Template Attacks in Principal Subspaces”, Cryptographic Hardware and Embedded Systems - CHES 2006, Lecture Notes in Computer Science Volume 4249, 2006, pp 1-14.
[6] Agrawal D, Rao J R, Rohatgi P, et al. Templates as master keys[C]// International Conference on Cryptographic Hardware and Embedded Systems. Springer-Verlag, 2005:15-29.
[7] Rechberger C, Oswald E. Practical Template Attacks[M]// Information Security Applications. DBLP, 2005:440-456.
[8] Arachambeau C, Peeters E, Standaert F X, et al. Template Attacks in Principal Subspaces[M]// Cryptographic Hardware and Embedded Systems - CHES 2006. 2006:1-14.
[9] Oswald D, Paar C. Breaking Mifare DESFire MF3ICD40: Power Analysis and Templates in the Real World[M]// Cryptographic Hardware and Embedded Systems–CHES 2011.Springer Berlin Heidelberg, 2011:207-222.
[10] Choudary O, Kuhn M G. Template Attacks on Different Devices[C]// International Workshop on Constructive Side-Channel Analysis and Secure Design. Springer International Publishing, 2014:179-198.
[11] Fan G, Zhou Y, Zhang H, et al. Towards optimal leakage exploitation rate in template attacks[J]. Security & Communication Networks, 2016, 9(16):3116-3126.
[12] L. Lerman, R. Poussier, G. Bontempi, O. Markowitch, and F. Standaert. Template attacks vs. machine learning revisited (and the curse of dimensionality in sidechannel analysis). In S. Mangard and A. Y. Poschmann, editors, Constructive Side-Channel Analysis and Secure Design - 6th International Workshop, COSADE 2015, Berlin, Germany, April 13-14, 2015. Revised Selected Papers, volume 9064 of Lecture Notes in Computer Science, pages 20–33. Springer, 2015.
[13] R. Gilmore, N. Hanley, and M. O’Neill. Neural network based attack on a masked implementation of aes. In Hardware Oriented Security and Trust (HOST), 2015 IEEE International Symposium on, pages 106–111, May 2015.
[14] C. Cortes and V. Vapnik. Support-vector networks. Mach. Learn., 20(3):273–297,Sept. 1995.49. B. Scholkopf and A. J. Smola.Learning with Kernels: Support Vector Machines,Regularization, Optimization, and Beyond. MIT Press, Cambridge, MA, USA,2001.
[15] A. Heuser and M. Zohner. Intelligent Machine Homicide - Breaking Cryptographic Devices Using Support Vector Machines. In W. Schindler and S. A. Huss, editors, COSADE, volume 7275 of LNCS, pages 249–264. Springer, 2012.

[16] Lerman L, Bontempi G, Markowitch O. Power analysis attack: an approach based on machine learning[M]. Inderscience Publishers, 2014.

[17] Maghrebi H, Portigliatti T, Prouff E. Breaking Cryptographic Implementations Using Deep Learning Techniques[M]// Security, Privacy, and Applied Cryptography Engineering. Springer International Publishing, 2016.