EDITE de Paris
École Doctorale Informatique, Télécommunications et Électronique

Résumé en français

CEA/CESTI-LETI

Extraction de Caractéristiques pour les Attaques par Canaux Auxiliaires

Eleonora Cagli
Id. 3373691

Directeur de Thèse
Emmanuel Prouff
Encadrante
Cécile Dumas

Laboratoire d'électronique et de technologie de l'information

Commissariat á l'énergie atomique et aux énergies alternatives
MINATEC Campus | 17 rue des Martyrs | 38054 Grenoble Cedex 9
www-leti.cea.fr
Établissement public á caractère industriel et commercial RCS Paris B 775 685 019

Direction de la recherche technologique

Contents

1 Contexte

1.1 Le CESTI

Les présents travaux de doctorat ont été réalisés au sein du laboratoire CESTI (Centre d'Évaluation de la Sécurité des Systèmes d'Information) du CEA de Grenoble. La mission d'un CESTI est d'évaluer les aspects sécuritaires des composantes embarqués qui nécessitent l'obtention d'un certificat pour pouvoir être commercialisés sur certains marchés sensibles. Les cartes à puces sont un exemple notable de tels types de dispositifs. Dans le schéma de certification franÃ§ais, c'est l'ANSSI (Agence National de la Sécurité des Systèmes d'Information) qui délivre le certificat, après consultation d'un rapport issu d'un des laboratoires CESTI agrées.

Un dispositif sécurisé permets, dans la grande majorité des cas, d'exécuter des algorithmes cryptographiques, pour offrir des garanties de confidentialité, authenticité, non-répudiation et intégrité des données pour les protocoles d'interface avec le dispositif-même. Quand un algorithme cryptographique est implémenté sur un support matérielle, il dévient potentiellement vulnérable à des attaques autres que ceux considéré en cryptanalyse classiques. En effet, outre à la faiblesse mathématique théorique de l'algorithme, des faiblesse matérielles liées à l'implémentation apparaissent. Ces attaques matérielles sont à prendre en compte dans une évaluation sécuritaire. Notamment, une partie du processus d'évaluation consistes à mener des attaques par canaux auxiliaires (ou *Side-Channel Attacks* en anglais, d'où l'acronyme SCA), qui font le sujet de cette thèse, et qui exploites des fuites d'information par des *canaux auxiliaires*, c'est-à-dire outre que les interfaces I/O du composant.

1.2 Les Attaques par Canaux Auxiliaires

Introduites en 1996 par Paul Kocher [4], les attaques par canaux auxiliaires sont basées sur l'observation des variations de certaines quantités physiques du composant, comme la consommation de puissance, ou le rayonnement électromagnétique, pendant l'exécution des algorithmes cryptographiques. En effet, en observant ces comportements physiques involontaires, qui viennent mesurÃ©s sous forme de signaux, des déductions sur les variables internes de l'algorithme peuvent être faites. Selon l'algorithme attaqué, faire inférence sur des variables internes bien choisies, les ceci-dites *variables sensibles*, est suffisant pour récupérer une clé secrète de l'algorithme.

2 Objectifs et Contributions

Dans un contexte d'évaluation d'un certain dispositif, les évaluateurs peuvent avoir accès à un ou plusieurs exemplaires du dispositif *ouverts*, ou *à secrets connus*. Ces dispositifs donnent droit à l'évaluateur de choisir ou connaître la clé secret cible d'une attaque, ou de fixer d'autres variables, de désactiver des contre-mesurer, ou de charger du logiciel. Cette possibilité est exploitée pour lancer des exécutions dans lesquelles l'attaquant aurait la connaissance complète du flux d'exécution, y compris les opérations, les variables internées manipulées, les access aux registres, les aléas tirer internement, ... En cette manière il est capable de comprendre et caractériser les relations entre le comportement interne du composant et les observations physiques, avant de lancer l'attaque. Quand une phase de caractérisation est disponible, on parle d'attaques *profilées*, qui ont un rôle très important dans l'évaluation d'un dispositif, permettant de testé celui-ci dans le scénario le plus favorable pour l'attaquant. Cette thèse se concentre principalement sur cette typologie d'attaques. En effet, nous traitons les problèmes qu'un évaluateur rencontre quand, dans un scénario si favorable, il veut exploiter de faÃ§on optimale la phase de caractérisation, pour extraire un maximum d'information des signaux acquis dans la phase propre d'attaque.Un de ces enjeux est la sélection des ceci-dits *Points d'Intérêt* (*Points of Interest* en anglais, ou PoIs), problème strictement relié au plus général problème de la réduction de dimension.

2.1 L'Avant-Propos de cette Thèse: la Recherche des Points d'Intérêt

L'acquisition des traces venant des canaux auxiliaires se fait habituellement à l'aide d'oscilloscopes numériques, qui effectuent un échantillonnage des signaux analogiques et les transforment en séquences numériques discrètes. Ces séquences sont souvent appelés *traces*, et leurs composants sont le *caractéristiques* temporelles, ou les points temporels, du signal. Pour garantir une inspection profonde

du dispositif, la fréquence d'échantillonnage doit être très élevée, ce qui provoque l'acquisition de traces de grand dimension. Cependant, il est attendu que seulement un nombre limité de points temporels soit relevant pour mener une attaque. Ce sont les PoIs, qui sont les points qui dépendent statistiquement de la variable sensible targette de l'attaque. En littérature l'utilisation de certains tests d'hypothèse statistique est déployée pour effectuer une sélection des PoI comme phase préliminaire d'une attaque. Cette sélection permettrait de réduire la complexité de l'attaque, en terme de temps et mémoire. L'objectif préliminaire de cette thèse était de proposer de nouvelles méthodes pour chercher et caractériser les PoIs, pour améliorer et possiblement optimiser ce pré-traitement des traces consistant en leur sélection.

2.2 Approche per Réduction de Dimension

Au-delà de l'utilisation de statistiques univariées pour identifier les PoIs, un différent axe de recherche s'est développé dans le contexte des SCAs, important du domaine de l'apprentissage automatique (ou *Machine Learning*, ML) des téchniques plus générales pour la réduction de la dimension des données, en passant d'une approche par sélection de caractéristiques à une approche par *extraction de caractéristiques*. Aux alentours du 2014, les méthodes linéaires d'extraction de caractéristiques ont attiré l'attention des chercheurs, en proposant l'application de techniques telles que l'*Analyse aux Composantes Principales* (PCA), l'*Analyse Discriminante Linéaire* (LDA) ou les *Projection Pursuits* (PP). Ces méthodes exploitent des combinaisons linéaires avantageuses des points temporelles des traces, pour définir des nouvelles caractéristiques amenant à des attaques plus efficaces. La première contribution de cette thèse fait partie de cette axe de recherche: on a abordé deux enjeux concertants l'application de PCA et LDA dans le contexte SCA: le choix des composantes, et le problème de la taille de l'échantillonnage. Les résultats de cette étude, publié en 2015 à CARDIS [1], sont résumés en Sec. 3.1 et font le sujet du Chapitre 4 de la thèse.

Aujourd'hui, tout dispositif demandant un certificat sécuritaire de haut niveau est équipé de contre-mesures spécifiques contre les SCAs. Une typologie de contre-mesure très efficace est le *masquage*. Quand un masquage est implémenté correctement, toute variable interne du calcul originaire qui est sensible, est divisée en plusieurs parties, dont la majorité est tiré au sort pendant l'exécution. Ceci est fait en sort que tout sous-ensemble propre des parties est statistiquement indépendant de la variable sensible elle-même. Le calcul cryptographique est mené an accédant uniquement aux parties, et non pas à la variable sensible. Ceci oblige l'attaquant à analyses des distributions de probabilité conjointes des caractéristiques signal, en étudiant conjointement son comportement aux instants temporels où chacune des parties est manipulée. Autrement dit, les statistiques univariées qui sont exploitable pour identifier les PoIs en absence de masquage deviennent inefficaces si un masquage est présent, car tout point temporel du signal est par lui-même indépendant de la variable sensible. En outre, les distribution jointes du signal doivent être analysées aux ordres statistiques supérieurs pour retrouver une dépendance statistiques des données sensibles. Ceci impliques que les méthodes linéaires d'extraction de caractéristiques sont aussi inefficace en ce contexte. Pour résumer, la sélection ou l'extraction de caractéristiques depuis traces protégées par masquage présente des difficultés non-négligeables. Cette complexité est mitigée quand l'attaquant peut effectuer une phase de caractérisation pendant laquelle il peut accéder aux valeurs aléatoires des parties du masquage pendant l'exécution. En pratique, ceci n'est pas tout le temps possible. Dans cette thèse on abord ce sujet dans le cas où cette possibilité est niée, en proposant l'exploitation de la technique de l'*Analyse Discriminante par Noyau* (*Kernel Discriminant Analysis*, KDA). Ceci est une extension de la LDA qui permet d'extraire des caractéristiques de faÃ§on non-linéaire. Les résultats obtenus dans ce contexte ont été publiés à CARDIS 2016 [2] et résumés en Sec. 3.2. Ils font le sujet du Chapitre 5 de la thèse.

2.3 Vers l'Apprentissage Profond

Si on observe le chemin qu'on a suivi pendant les travaux de thèse, on remarque qu'on est partie du problème d'identifier les PoIs d'un signal, ce qui est classiquement résolu par des outils statistiques classiques, ensuite on a élargi à la fois les objectifs et les méthodologies. En effet, que ce qui plus influenÃ§ait la réussite d'une attaque était la qualité de l'extraction d'information. Extraire de l'information demande d'approximer des distributions de probabilité qui permettent de distinguer différent valeurs secrètes. Les premiers attaques par canaux auxiliaires proposées en littérature opérait point par point, donc nécessitait d'analyser les distributions de donnée en quelques instants temporels pris singulière-

ment. Dans ce contexte la sélection des PoIs jouait un rôle fondamental. Cependant, dès qu'on fait un pas en arrière vers l'objectif d'une attaque, et qu'on se demande comment approximer des distributions distinguables, le fait de rejeter complètement une grande partie des caractéristiques du signal, en en sélectionnant que quelques unes, paraÃ®t du gaspillage. Des méthodes appropriées pour combiner ces caractéristiques peuvent mener à l'extraction de caractéristiques plus discriminantes. Pour déterminer ces combinaisons appropriées, nous avons exploré les outils d'extraction de caractéristiques afin de les utiliser comme pré-traitement du signal. En un premier temps, nous avons considéré des outils linéaires, ensuite des généralisations non-linéaires pour satisfaire une condition nécessaire à adresser les implémentations protégées par masquage.

Conscients du fait que ces outils sont à mi-chemin entre les statistiques multivariées classiques et le domaine de l'apprentissage automatique, nous avons commencé à explorer ce domaine, qui est aujourd'hui en grand développent. Le grand intérÃªt attiré par l'apprentissage automatique est justifié de la tendance à capter et analyser données de grand dimension dans une large variété de champs applicatifs, y compris les attaques par canaux auxiliaires. Pour cela, des modèles de plus en plus complexes sont mis en oeuvre, trop complexes pour Ãªtre traités dans un cadre de statistiques formelles. L'apprentissage automatique accepte des non-optimalité intrinsèques mais fait démonstration aujourd'hui d'excellents résultats.

L'étude des outils d'apprentissage automatique nous a mené a effectuer davantage un pas en arrière vers l'objectif d'une attaque: plutôt que optimiser des pré-traitement de données, afin d'obtenir des caractéristiques montrant des distributions facilement distinguables, nous pouvons chercher des modèles pour approximer directement ces distributions à partir des données brutes. Cette approche est propre d'une branche de l'apprentissage automatique, qui s'appelle apprentissage profond. Dans l'apprentissage profond la phase de caractérisation des données est effectuées en un seul processus, qui intègre éventuellement les pré-traitements nécessaires. Ceci est fait à l'aide de modèles multicouches, notamment les *réseaux neuronals* (*Neural Networks*, NN), sur lesquels on se concentre dans la dernière partie de la thèse. étant des modèles non-linéaires, les NN peuvent Ãªtre utilisés pour adresser la contremesure de masquage. De plus, des architectures particulières de NN, les ceci-dits réseaux convolutifs (CNN), conÃ§us originairement pour la reconnaissance d'image, s'adapte aussi bien à d'autres types de contremesures: celles qui provoquent de la désynchronisation des signaux. Nous avons étudié ce contexte, en proposant l'utilisation des CNNs comme solution, équipés d'une autre stratégie classique dans le domaine de l'apprentissage automatique, l'*augmentation des données* (DA). Le Chapitre 6 de la thèse est dédié à ce sujet. Les résultats obtenus ont été publiés à CHES 2017 [3] et résumés en Sec. 3.3.

3 RÃ©sultats principaux

3.1 Techniques LinÃ©aires de RÃ©duction de Dimension

3.1.1 LDA and the Small Sample Size problem

3.2 Analyse Discriminante par Noyau

3.3 RÃ©seau Neuronal Convolutif

4 Conclusions et Perspectives

References

[1] Eleonora Cagli, Cécile Dumas, and Emmanuel Prouff. Enhancing dimensionality reduction methods for side-channel attacks. In *International Conference on Smart Card Research and Advanced Applications*, pages 15–33. Springer, 2015.

[2] Eleonora Cagli, Cécile Dumas, and Emmanuel Prouff. Kernel discriminant analysis for information extraction in the presence of masking. In *International Conference on Smart Card Research and*

Advanced Applications, pages 1–22. Springer, 2016.

[3] Eleonora Cagli, Cécile Dumas, and Emmanuel Prouff. Convolutional neural networks with data augmentation against jitter-based countermeasures - profiling attacks without pre-processing. In Wieland Fischer and Naofumi Homma, editors, *Cryptographic Hardware and Embedded Systems - CHES 2017 - 19th International Conference, Taipei, Taiwan, September 25-28, 2017, Proceedings*, volume 10529 of *Lecture Notes in Computer Science*, pages 45–68. Springer, 2017.

[4] Paul C Kocher. Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems. In *Annual International Cryptology Conference*, pages 104–113. Springer, 1996.