Cryptanalyse par side channel

J-J. Quisquater & D. Samyde

Université catholique de Louvain
Groupe Crypto
3 Place du Levant
B-1348 Louvain la Neuve, Belgium
jjq@dice.ucl.ac.be
samyde@dice.ucl.ac.be

Résumé

La cryptologie rassemble l'ensemble des techniques de cryptographie et de cryptanalyse. La cryptographie respecte les principes de Kerckhoffs, qui justifient que toute information liée à un cryptosystème peut être publique à l'exception des clés de chiffrement. La cryptanalyse inclut donc des techniques très avancées afin de retrouver ces clefs. La controverse sur la sécurité de l'algorithme de chiffrement DES aura grandement contribué au développement et à l'avènement de nouvelles méthodes mathématiques de cryptanalyse. Les attaques linéaires et différentielles en sont les exemples les plus probants. Toutefois bien que ces techniques nécessitent encore souvent de grandes quantités de paires de textes en clairs et de textes chiffrés, il existe d'autres méthodes très puissantes basées sur les "fuites d'information" involontaires. En effet un cryptosystème peut laisser fuir de l'information de différentes manières, c'est ainsi que des données sensibles peuvent parfois être extraites de signaux physiques émis par une machine de chiffrement. La température, les ondes acoustiques, le rayonnement électromagnétique, le temps de calcul ou la lumière (infra rouge, rayonnée, interaction avec un laser) sont autant d'indices qui peuvent s'avérer extrêmement dangereux. On parle alors de side channel. La cryptanalyse par les side channel a longtemps été le terrain de compétence réservé des services secrets, mais depuis une dizaine d'années, les mondes scientifique et universitaire contribuent à développer des nouvelles techniques de side channel très efficaces.

Keywords Side channel cryptanalysis, Tempest, Timing attack, Power & ElectroMagnetic Analysis, Fault Analysis

Introduction

L'histoire de la cryptanalyse conventionnelle et celle du développement des ordinateurs les plus puissants sont intimement liées. Il existe toutefois bien des anecdotes très impressionnantes quant à la puissance des techniques mises en oeuvre pour une époque donnée. La preuve incontestable en est la brillante cryptanalyse de l'Enigma marine par A. Turing et les deux Colossus du GHCQ à Bletchley Park pendant la deuxième guerre mondiale. Comme pour les techniques de cryptanalyse conventionnelles, la puissance des side channels étant redoutable, leurs réussites sont longtemps demeurées dans l'ombre.

Il y a quelques siècles, les automates les plus avancés reposaient sur des principes mécaniques aujourd'hui bien maîtrisés. Avec l'évolution de la science, l'électricité à pris le pas sur la mécanique et ce fut l'avènement de l'électromécanique puis de l'électronique. Toutefois bien que les effets physiques désirés soient utilisés à bon escient, et constituent la composante principale du fonctionnement d'une machine donnée, d'autres effets parasites ont parfois été négligés. Dans ce cas, il devient alors parfois possible de les utiliser pour extraire les informations sensibles manipulées par un appareil [1].

Lorsqu'un courant parcourt un conducteur, il crée un champ électrique, un champ magnétique ainsi qu'un échauffement. Ces effets sont connus et décrits par les équations de Maxwell et la loi de Joules. Mais lorsque certains de ces effets physiques existants ne sont pas parmi les volontés premières du concepteur d'un automate, ils peuvent trahir son fonctionnement. Il est possible par exemple de capter le champ électromagnétique à distance avec une antenne appropriée, de le réamplifier, de démoduler pour obtenir des informations. Dans le cas d'un tube cathodique de télévision ou d'un écran d'ordinateur, le pixel à l'écran est obtenu grâce à la projection d'un faisceau d'électrons sur des molécules sensibles. Mais la position du faisceau d'électrons est contrôlée par des bobines de déplacement. Le champ qu'elles rayonnent si il est capté, amplifié puis représenté à l'entrée de d'autres bobines permet de copier l'image de l'écran à distance.

Pour ce qui concerne la chaleur, certaines caméras thermiques autorisent une assez bonne résolution, permettant la construction de modèles de signatures. Ces modèles caractérisent alors la machine. Ces techniques sont très utilisées par certains systèmes d'armes récents, mais les fabricants de cryptoprocesseurs s'ingénuent également à essayer de limiter les échauffements caractéristiques révélant trop d'informations sur le calcul en cours dans leur puce. La liste exhaustive des effets de bord indésirables lors du fonctionnement d'un appareil ne saurait être donnée ici, toutefois cet article s'emploie à rappeler l'importance de la cryptanalyse par les side channel et à dresser une liste des attaques les plus utilisées actuellement contre les implémentations cryptographiques.

Souvent on oublie de souligner l'important rôle joué par les side channel dans des cas de cryptanalyses réussies. Il existe plusieurs manières de concevoir la cryptanalyse. La première et la plus conventionnelle consiste à regarder les primitives cryptographiques comme des objets mathématiques ou des algorithmes. L'autre conception s'intéresse plutôt à l'implémentation de la primitive dans une machine de chiffrement. Il est alors intéressant d'analyser les effets de bord obtenus. Un source de side channel est donc inhérente à la structure même de l'implémentation physique et elle peut grandement faciliter le travail d'un assaillant éventuel.

Historique

Il est bien difficile de fixer avec précision la naissance de la cryptanalyse par side channel. Mais il serait faux de fixer cette date à la fin du XXème siècle. Il semble même que l'on puisse fixer cette date à la fin du XIXème siècle ou au tout début du XXème. Les premières prises de conscience de l'existence de side channel n'ont visiblement pas donné lieu à des publications et à des communications internationales.

J.Maxwell établit sa théorie sur les ondes électromagnétiques en 1873. Mais dès la fin du 19ème siècle, des problèmes de crosstalk apparaissent dans les liaisons téléphoniques. Pendant la première guerre mondiale ces problèmes de couplage électriques furent utilisés pour espionner des communications. Les informations obtenues étaient uniquement recopiées sur un autre média et écoutées. Le traitement du signal étaient inexistant dans de telles interceptions.

Mais en 1918 H. Yardley et son équipe découvrirent que des informations classifiées pouvaient s'échapper de matériels électriques, et que ces fuites permettaient de retrouver les secrets manipulés. Les données contenues dans les appareils de chiffrement modulaient un signal sur la bande d'une source d'enregistrement proche. Au milieu des années trente, les études sur la machine à écrire d'IBM indiquèrent que les fuites d'informations étaient importantes et devaient être prises en considération.

Les militaires prirent alors au sérieux ce genre de fuites et les différentes armées occidentales se mirent à faire très attention à limiter les rayonnements sur leurs appareils sensibles lors de la deuxième guerre mondiale, particulièrement pour les oscillateurs embarqués. Par la suite cela n'empêcha pas les télétypes communistes d'être "interceptés" à Berlin. Les antennes de réception étaient placées dans des tunnels proches. Durant les années cinquante, les autorités chinoises utilisèrent également des techniques acoustiques pour espionner des ambassades, alors que les russes envoyaient des jets de micro-ondes sur des barres de métal contenues dans des statues, afin d'écouter les ondes acoustiques d'une ambassade américaine. La création de la NSA en 1953 donna immédiatement une importance cruciale à la récupération des signaux compromettants. Le mot SIGINT qui est l'acronyme de SIGnal INTelligence prit alors tout son sens. Une des premières interception de rayonnement électromagnétique connue fut réalisée en utilisant les fils téléphoniques comme antenne. Par le futur, ce genre de technique continuera d'être appliquée avec succès pour des interceptions d'écrans informatiques à plusieurs centaines de kilomètres de distance. Ces premières expériences furent réalisées au M.I.T.

Les années cinquante fut le début de la grande série des normes militaires américaines & OTAN, en ce qui concerne la limitation des rayonnements électromagnétiques compromettants et l'utilisation de blindage électromagnétiques. Les militaires américains s'inquiétèrent de cette nouvelle menace de compromission et initièrent alors le programme TEMPEST. Au début des années soixante, l'ambassade russe et certains appareils francais à Londres furent espionnés grâce à la récupération de leurs rayonnement électromagnétiques. En démodulant correctement les signaux récupérés, les anglais arrivaient à retrouver des informations classifiées. Cette fois, le traitement du signal devint indispensable et prit un rôle de plus en plus grand dans le futur [Figure 1].

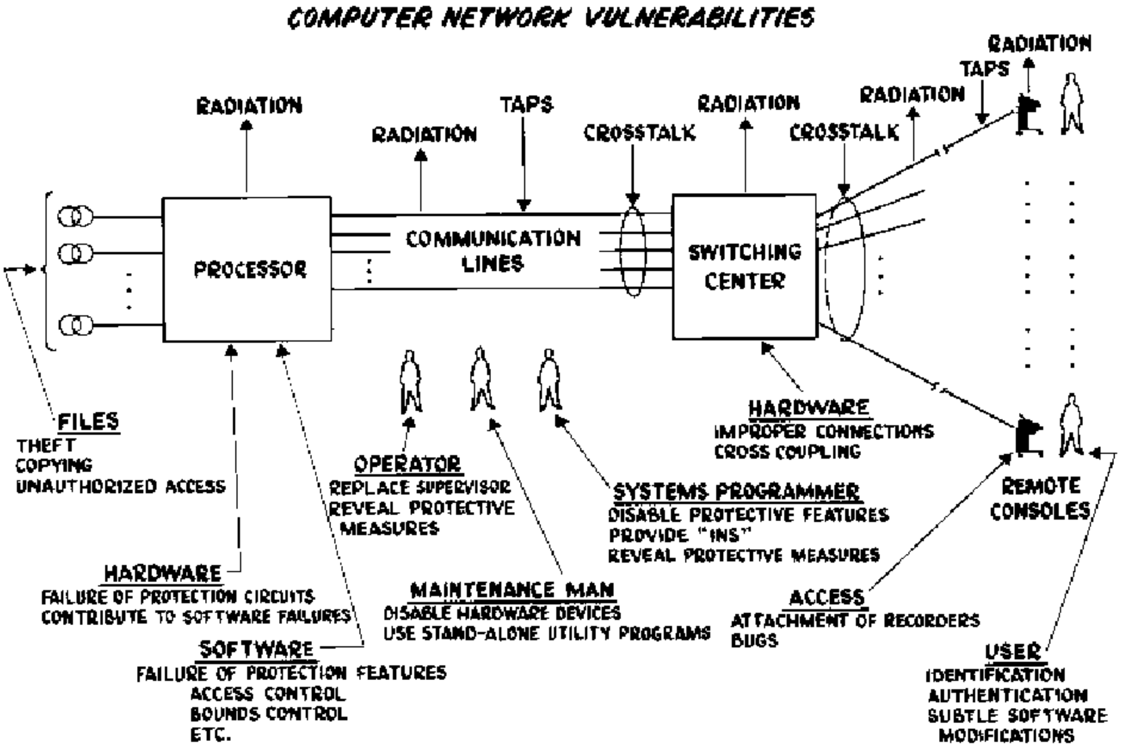

Figure 1 : Les sources de fuites par un des concepteurs d'Arpanet

Un des exemples les plus frappants concerne la cryptanalyse d'un chiffrement russe lors de la crise de Cuba. Le navire américain intercepteur Oxford réussit alors à capter les radiations électromagnétiques émises par une machine de chiffrement soviétique située sur l'île. De plus la marine américaine réalisa alors que la mesure du bruit rayonné permettait de connaître la position des rotors à l'intérieur de certaines machines de chiffrement.

Les civils n'étaient alors pas très au courant des possibilités d'interception, car ce genre de méthodes étaient classifié. Mais rapidement, dès les années soixante dix, des particuliers mentionnèrent des cas d'interférences entre différents matériels électroniques. Et les cas de jeunes chercheurs faisant jouer un air de musique sur un poste de radio ou un récepteur munit d'une démodulation, en utilisant les commutations d'un écran ou d'une imprimante ne sont pas isolés.

Toutefois l'utilisation des analyses par side channel continuèrent d'exister et l'opération Kilderkin, qui consistait dans la récupération des émanations électromagnétiques de l'ambassade russe au Canada, prit place au milieu des années soixante dix. A peu près à la même époque les services de renseignements polonais furent pris en flagrant délit d'espionnage sur du matériel militaire soviétique ; une fois de plus les rayonnements électromagnétiques furent mis en cause. Toutefois cette fois-ci, c"etaient les fils d'alimentation électrique qui étaient écoutés. En 1984 la NSA introduit de manière non classifiée le concept de zone de sécurité autour d'un point de mesure et c'est l'année suivante que IBM, bien après Zenith, construit son premier PC blindé contre les fuites électromagnétiques [2]. A la fin des années quatre vingt I. Murphy, aussi connu sous le nom de Captain Zap, publia les premiers plans de récepteurs TEMPEST, après que la télévision anglaise ait montré des démonstrations à l'antenne. Dans les années quatre vingt, l'armée francaise s'intéressa spécifiquement au développement d'une chaîne de mesure TEMPEST. Les services de renseignement bulgares utilisèrent des camions espions équipés d'antennes et maquillés pour intercepter des communications militaires dans les pays occidentaux. Dans les années quatre vingt dix, R. Anderson et M. Kuhn travaillŕent sur les interceptions de rayonnements et publiŕent des polices permettant de réduire les fuites. De toute manière, les normes de santé liées à la puissance des émissions et la prise de conscience des constructeurs firent que les niveaux de fuites se firent de plus en plus faibles. En 1996 P. Kocher publia ses travaux basés sur la mesure du temps de calcul d'un cryptosystème. La timing attack exploitée par F. Koeune permit de récupérer une clé privée RSA de 512 bits avec quelques centaines de milliers d'échantillons en quelques minutes alors que la factorisation d'un nombre équivalent requiérait plusieurs mois de calculs intensifs. P. Kocher et de son équipe se concentrèrent alors sur l'analyse de consommation, et ils mirent en évidence la puissance des mesures différentielles. C'est ainsi que les fondeurs de silicium se vient demander par leurs gros clients bancaires de trouver des solutions rapides et efficaces à ce problème. Les années quatre vingt dix viren apparaître un autre type de cryptanalyse, basé sur l 'injection de fautes. Actuellement la mesure des rayonnements électromagnétiques des processeurs cryptographiques en champ proche, couplée à une mesure différentielle permet de mettre à mal bien des implémentations. Les progrès des contre mesures réalisés dans le domaine de la maîtrise des effets de l'insertion de faute sont eux quasiment décisifs.

L'histoire relate même des cas ou les side channel furent introduits volontairement afin de pouvoir faciliter la compromission des données sensibles par la suite. Au final l'utilisation des side channel s'est développée de plus en plus pour devenir aujourd'hui une des composantes essentielles de la collecte de renseignement.

Les fuites par rayonnements électromagnétiques

Les fuites par rayonnements électromagnétiques sont encore exploitables de nos jours contre des systèmes récents. Elles peuvent être utilisées pour compromettre tous types d'appareils.

La qualité des antennes de réception est très importante, et il en existe de plusieurs types (logarithmique, fouet, planaires...). De même, une excellente stabilité en fréquence des oscillateurs locaux assure un bon résultat ; les meilleurs récepteurs autorisent une précision de 10^{-6} Hz en utilisant des boucles à verrouillage de phase imbriquées.

Les contre mesures s'étant développées avec le temps, dans le cas d'un tube cathodique, il est difficile d'obtenir une interception valide à plus d'une dizaine de mètres, avec du matériel ancien. Cette distance et ce niveau de rayonnement sont à comparer avec les distances obtenues par les satellites espions qui, de par leur position alignée avec une ligne d'antennes terrestre, écoutent certaines transmissions depuis l'espace.

De nos jours, beaucoup d'informations sont accessibles parce qu'elles ont été déclassifiées ou ont percolé dans le public. Mais la simple présence ou non de rayonnement électromagnétique permet bien souvent de fournir une information utile à un assaillant. Quelques systèmes vont jusqu'à recréer un faux champ magnétique autour d'eux de manière à masquer leur présence ou leurs rayonnements. Toutefois, le théorème de C. Shannon indique qu'en réitérant la mesure, il devient alors possible de supprimer le bruit et d'obtenir un rapport signal sur bruit aussi élevé que souhaité.

Des appareils comme les cartes à puces utilisent des contremesures opératives pour détecter la répétition d'une exécution particulière un grand nombre de fois. Dans la grande majorité des cas, les concepteurs se contentent d'essayer de limiter le niveau de rayonnement ou alors de définir une zone de protection ne permettant pas d'obtenir un signal intelligible. Mais des applications nécessitent bien souvent d'être protégées par une cage de Faraday, et il n'est pas toujours simple et possible d'en utiliser une.

Certaines recommendations liées au niveau d'atténuation des cages de Faraday semblaient inexplicables par le passé ; les récents travaux de M. Kuhn et d'autres universitaires semblent pouvoir expliquer les différences entre les atténuations requises par les militaires et les civils. En effet les atténuations militaires ne semblent laisser aucune chance d'interception même avec du matériel récent, alors que les niveaux civils n'assuraient pas cette qualité de blindage.

Il est également important de noter que la lumière issue d'un écran contient une information utile qui n'est autre que le signal vidéo. Comme l'a récemment publié M. Kuhn, il est possible de reconstruire une image vidéo à partir de la luminosité d'un écran distant.

Les fuites par mesure du temps

Une des premières attaques, par mesure du temps de réponse d'un système, permettait de réduire le nombre d'essais pour retrouver un mot de passe. Cette attaque était particulièrement efficace contre les implémentations d'Unix. En effet, si la comparaison entre le mot de passe enregistré et le mot de passe saisi était effectuée octet par octet en partant du premier bit, et que la réponse de la machine était disponible dès la première erreur, il était alors simple de tester tous les octets possibles en première position et de choisir celui qui avait le temps de réponse le plus long. En effet, ce dernier était forcément identique à l'octet recherché. L'augmentation du temps de réponse correspondait à la somme du temps de la comparaison valide du premier octect et du temps de réponse erronée du deuxième. Selon R. Moreno, ce genre d'analyse a également fonctionné lors du portage d'une application bancaire depuis un processeur Motorola sur une carte à puce incluant un processeur Thomson. L'application de cette attaque a été réalisé par F. Grieu. Il est également possible d'utiliser la mesure du temps de réponse pour connaître le contenu du cache d'un serveur informatique distant. En effet, une information déjà présente dans le cache de la machine ne nécessitera pas le même temps de réponse que si l'information doit être chargée.

En 1996 P. Kocher publia un article assez théorique sur l'utilisation de la mesure de temps, pour essayer de retrouver des clés cryptographiques. Deux ans plus tard, F. Koeune mis en application cette analyse et démontra dans une publication comment il était possible de mettre à mal une implémentation naive de l'algorithme Square & Multiply servant à l'exponentiation modulaire. Selon la valeur du bit de clé traité, il n'y a que deux possibilités de branchement à l'intérieur du code exécuté pour sa démonstration. Mais ces deux branchements ne prenant pas le même temps d'exécution, il est alors possible de connaître les clés cryptographiques assez rapidement [3], avec quelques centaines de milliers d'échantillons.

Les contre mesures à ce genre d'analyse peuvent paraître triviales, mais une réponse à temps constant n'est pas toujours la mieux appropriée : il faut quelques fois considérer le pire cas d'un algorithme, ou insérer des contre mesures plus subtiles, modifiant les propriétés temporelles.

Les fuites par mesure de la température

Les analyses en température sont peu utilisées contre les implémentations cryptographiques ; elles sont plus courantes dans l'analyse d'images aériennes ou spatiales. Dans bien des cas, des matériels stationnés à un endroit donné ont modifié la température ou l'illumination de leur lieu de stockage. Même longtemps après leur départ il est toujours possible de procéder à une mesure révélant leur présence passée.

Dans le cas des processeurs, le facteur limitant est la diffusion de la chaleur. En effet, l'équation de propagation est bien connue, mais les temps de propagation sont rédhibitoires. Toutefois, il est important de ne pas obtenir de points chauds sur la puce, afin de ne pas révéler une activité spécifique [4]. Par le passé, les testeurs de composants utilisaient des cristaux liquides pour révéler les zones d'activité du composant. Cette méthode ne possède pas une dynamique suffisante et se révèle trop lente pour être réellement puissante.

Les fuites par mesure de la consommation

Une petite antenne insérée dans un condensateur volontairement troué, peut fournir beaucoup d'informations sur l'activité d'une carte électronique. Cette ancienne technique est bien connue et utilisée depuis longtemps. Une carte électronique ne contient souvent que peu d'éléments très intéressants à analyser. Il suffit alors d'essayer de mesurer leur consommation. En insérant une résistance de faible valeur entre la broche de masse et la masse générale de la carte, il est alors possible de traquer la consommation spécifique d'une puce. Par le passé, quelques personnes avaient déjà remarqué que l'analyse de la consommation d'un cryptoprocesseur pouvait fournir de l'information sur les clés manipulées. P. Kocher [5], en 1998, a introduit la notion de mesure différentielle ; il s'agit de l'intercorrélation de deux variables aléatoires. Cette technique appelée DPA s'avère être très dangeureuse contre les cartes à puces et les puces les plus récentes. De plus, il est possible d'améliorer encore les résultats, en remplacant la résistance de mesure par une bobine située dans le champ proche du processeur. L'amélioration du rapport signal sur bruit est alors comprise entre 30 et $40dB$; on parle alors d'analyse électromagnétique [6,7]. Bien que cette technique date du début de l'an 2000, les fabricants de composants commencent à la prendre au sérieux car elle semble menacante. L'algorithme utilisé est simple dans les deux cas : il suffit de quantifier l'influence d'un bit de la clé et de séparer les traces obtenues en deux ensembles. En vérifiant ensuite l'hypothèse sur la valeur du bit de travail, il est possible de connaître la valeur du bit de clé recherchée.

Conclusion

En conclusion, la cryptanalyse utilisant les side channel est puissante et peut rendre inefficace des implémentations d'algorithmes très robustes et bien construits. Dans le futur, d'autres side channels feront peut-être leur apparition, mais il semble toutefois que le coût réel des ces attaques soit de plus en plus élevé. Afin de se préserver d'un bon nombre de cryptanalyses, les implémentations doivent maintenant intégrer un niveau d'expertise non négligeable. Les contremesures sont toujours possibles et disponibles, mais elles doivent être bien pensées. Il est facile d'espérer éviter un side channel et en fait de se fragiliser vis à vis d'un autre [8,9].

Dans cet éternel jeu du gendarme et du voleur, les cryptanalystes semblent pour le moment avoir la part belle [10] ; le futur fera sans doute évoluer les choses rapidement.

Références

[1] NACSIM 5000 : *Tempest Fundamentals*, National Security Agency, Fort George G.Meade, Maryland. Feb 1982. Partially declassified also available at `http://cryptome.org/nacsim-5000.htm`.

[2] M. Kuhn and R. Anderson, Soft tempest : *Hidden data transmission using electromagnetic emanations*, In D. Aucsmith, editor, Information Hiding, vol 1525 of Lecture Notes in Computer Science, pp 124-142. Springer-Verlag, 1998.

[3] J. Kelsey, B. Schneier, D. Wagner, and C. Hall, *Side Channel Cryptanalysis of Product Ciphers*, in Proc. of ESORICS'98, Springer-Verlag, September 1998, pp. 97-110.

[4] J-S. Coron, P. Kocher, and D. Naccache, *Statistics and Secret Leakage*, Financial Cryptography 2000 (FC'00), Lecture Notes in Computer Science, Springer-Verlag.

[5] P. Kocher, J. Jaffe and B. Jun, *Differential Power Analysis*, In M. Wiener, editor, Advances in Cryptology - CRYPTO'99, vol. 1666 of Lecture Notes in Computer Science, pp. 388-397, Springer-Verlag, 1999. Also available at `http://www.cryptography.com/dpa/Dpa.pdf`.

[6] K. Gandolfi, C. Mourtel and F. Olivier, *Electromagnetic analysis : concrete results*, In Koç, Naccache, Paar editors, Cryptographic Hardware and Embedded Systems, vol. 2162 of Lecture Notes in Computer Science, pp. 251-261, Springer-Verlag, 2001.

[7] J.-J. Quisquater and D. Samyde, *ElectroMagnetic Analysis (EMA) Measures and Counter-Measures for Smart Cards*, in I. Attali and T. Jensen, editors, E-Smart Smartcard Programming and Security, vol. 2140 of Lecture Notes in Computer Science, pp. 200-210, Springer-Verlag 2001.

[8] R. Anderson, M.Kuhn, *Tamper Resistance - A Cautionary Note*, Proc. of the Second USENIX Workshop on Electronic Commerce, USENIX Association, 1996.

[9] O. Kommerling and M. Kuhn, *Design principles for tamper-resistant smartcard processors*, In Proc. of the USENIX Workshop on Smartcard Technology (Smartcard'99), pp. 9–20. USENIX Association, 1999.

[10] E. Biham and A. Shamir, *Power Analysis of the Key Scheduling of the AES Canditates*, in Second Advanced Encryption Standard Candidate Conference, Rome, March 1999.