Histoire de la carte à puce du point de vue d'un cryptologue

Louis Guillou

Expert émérite
Division R&D de France Telecom
MAPS/DPC, 4 Rue du Clos Courtel,
BP 91226, 35512 Cesson Sévigné, France
Tél 02 9912 4247 Fax 02 9912 3600
louis.guillou@fancetelecom.com

Résumé. Il y a concomitance entre les débuts de la carte à puce et les premiers pas de la cryptologie dans le domaine public. Aujourd'hui, sans cryptographie appropriée, la carte à puce ne conviendrait ni pour les banques, ni pour la télévision à péage, ni pour le téléphone mobile, ni pour la santé, et ainsi de suite. Le lien entre carte à puce et cryptologie est très fort : la carte confine des clés et des algorithmes ; elle contrôle son propre usage ; elle reconnaît son porteur. Bien sûr, la sécurité absolue n'existe pas, mais la sécurité peut toujours s'améliorer. La sécurité des cartes repose sur des logiciels spécifiques, évalués selon la méthodologie des critères communs et des profils de protection.

Abstract. *The start of smart card coincides with the advent of cryptology in the public domain. Today, without an appropriate cryptography, the smart card would be inappropriate for banking, pay-TV, mobile phone, health, and so on. The link between smart cards and cryptology is very strong: the smart card confines keys and algorithms; it controls its own use; it recognizes its holder. Absolute security does not exist, but security may always be improved. Card security relies on specific software evaluated according to common criteria methodology and protection profiles.*

1 Les débuts de la carte à puce

1.1 Les premiers brevets

Les développements de produits avancés ne sont jamais le fruit des idées d'un seul homme, surtout si ce dernier ne dispose pas de la technologie nécessaire. Jules Verne inventa-t-il la fusée pour aller dans la lune ? Ne fallut-il pas attendre Von Braun et bien d'autres ? En fait, les débuts de la carte à puce ressemblent à ceux de l'aviation : beaucoup rêvaient de voler sur de drôles de machines sans y parvenir.

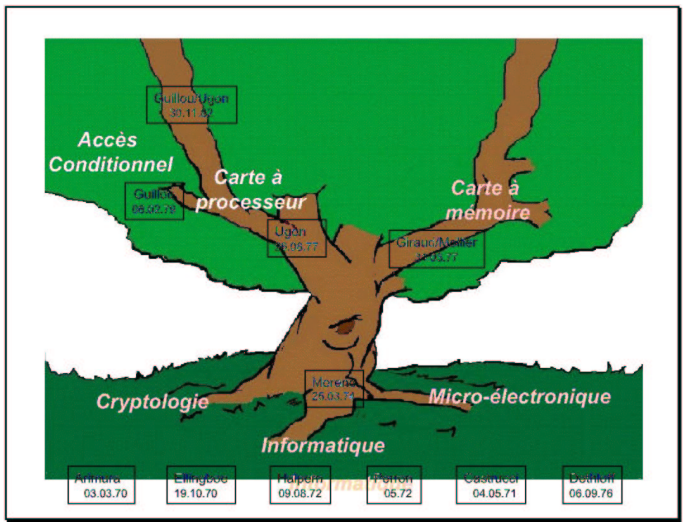

FIGURE 1 – Généalogie des premiers brevets de la carte à puce

Comme en témoignent les nombreux brevets « pionniers » déposés au fil du temps, l'utilisation d'un composant électronique dans un objet au format d'une carte de crédit a fait l'objet de réflexions au Japon, aux États-Unis et en Europe. En 1967, le japonais Arimura introduit un algorithme d'authentification dans une carte. En 1970, l´américain Ellingboe montre une carte dotée d'un registre sériel d'identification piloté par un processeur. En 1973, le « stylo électronique » de Halpern anticipe toutes les fonctions de sécurité d'une carte prépayée sans contact pour le paiement des bus à San Francisco. Si les brevets « pionniers » ne donnent pas lieu à réalisation, s'ils ne « germent » pas, c'est parce que le terrain n'est pas propice, que la technologie nécessaire, à savoir, la microélectronique, l'informatique et la cryptologie, n'a pas encore atteint un stade suffisant de développement.

En 1974, les brevets « fondateurs » de Moreno décrivent un objet portable à mémoire revendiquant « des moyens inhibiteurs », « un comparateur avec un compteur d'erreurs » et « des moyens de couplage avec le monde extérieur ». Il convient de souligner le rôle essentiel du rédacteur de ces brevets, Jean Moulin.

En 1977, Ugon pour la compagnie CII HB (CII Honeywell Bull) jette les bases de la carte à processeur, et comme cette carte permet d'exécuter des algorithmes cryptographiques, dès 1978, pour le CCETT (Centre Commun d'Etudes de Télédiffusion et Télécommunications), j'y pense comme dispositif d'accès aux services audiovisuels. Grâce à cette nouvelle dimension, la carte est prête à conquérir tous les marchés de la sécurité. Â partir des brevets « fondateurs » de Moreno, sous l'impulsion de la compagnie CII HB et du CCETT, c'est tout un arbre qui se développe, illustré

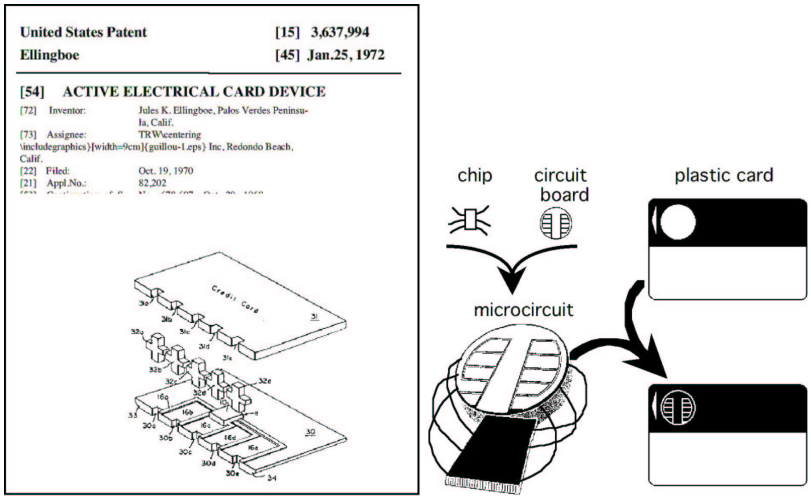

FIGURE 2 – Comparaison de deux procédés de fabrication

par la figure 1, avec deux ramifications principales : la carte à processeur et la carte à mémoire.

1.2 Les premiers développements

La première licence. Dès 1973, la compagnie CII HB s'intéresse de près au brevet pionnier de Halpern, puis acquiert en 1974 une licence des brevets fondateurs de Moreno.

En 1975, la compagnie CII HB crée une division dotée d'importants moyens de recherche. Ugon assure dès le début l´animation technique de cette division avec pour mission de sécuriser les systèmes d'information. Afin de convaincre les utilisateurs potentiels de commencer des expérimentations, les recherches portent sur tous les éléments nécessaires à la mise en place de systèmes utilisant des cartes à puce. La figure 2 illustre le rêve et la réalité en comparant deux procédés de fabrication.
– Le procédé de fabrication décrit dans le brevet pionnier d'Ellingboe n'a jamais été utilisé.
– Le procédé de fabrication mis au point par CII HB est largement utilisé à partir de 1979.

La carte à deux puces. Le premier objet fonctionnel à puce au format d'une carte de crédit apparaît le 21 mars 1979, fruit d'une étroite coopération avec la société Motorola. Appelée CP8, cette carte établit la validité du concept, c'est-à-dire, la coexistence d'un microcircuit et de ses contacts avec une zone d'embossage et

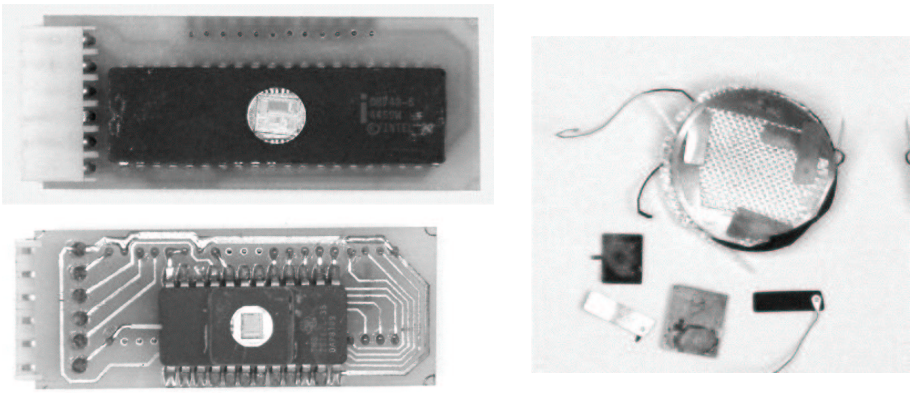

FIGURE 3 – Un dispositif à deux puces et les restes de la première carte à une seule puce

des pistes magnétiques sur une carte aux normes ISO/CEI 7810 et 7811. L'objet porte deux composants largement disponibles sur le marché : un microprocesseur masqué et une mémoire programmable électriquement, à savoir, un composant CPU 3870 conçu à l´origine par la société Fairchild, et un composant EPROM 2716, effaçable par rayonnement ultraviolet, conçu par la société Intel. En 1981, avec une carte à deux puces, la compagnie CII HB expérimente le télépaiement à domicile à Vélizy en collaboration avec la Poste (c'était du commerce électronique avant l'heure).

Pour les démonstrations d'accès conditionnel aux magazines diffusés par le télétexte Antiope [10], dès 1978, j'utilise un ordinateur en kit (Intel SDK, environ six kilogrammes sans l´alimentation), puis en 1979, un dispositif comprenant un connecteur à six broches, un processeur et une mémoire, tous deux programmables électriquement, à savoir deux composants Intel 8748 et 2716. Plutôt que la faisabilité d'une carte, déjà établie par CII HB, mon but est d'établir la faisabilité d'un dispositif cryptographique portant et gérant les droits et les clés de l'usager. Convaincant pour un public techniquement averti, le dispositif en figure 3 est quelque peu perturbant pour un public peu averti : ce n'est pas une carte.

La solution naturelle consiste à utiliser une seule puce dans la carte, pour au moins trois raisons : coût, fiabilité et sécurité du produit final. En effet, la fabrication des cartes est plus simple ; le risque de panne est réduit ; et enfin, il n´y a pas de connexions d´une puce à une autre, ce qui évite un accès facile entre le processeur et la mémoire programmable électriquement.

La toute première carte à une seule puce. Au cours de l´été 1980, la Direction Générale de Télédiffusion de France envisage une démonstration d´accès conditionnel au plus haut niveau de l´État. Pour démontrer un usage cryptographique sans ges-

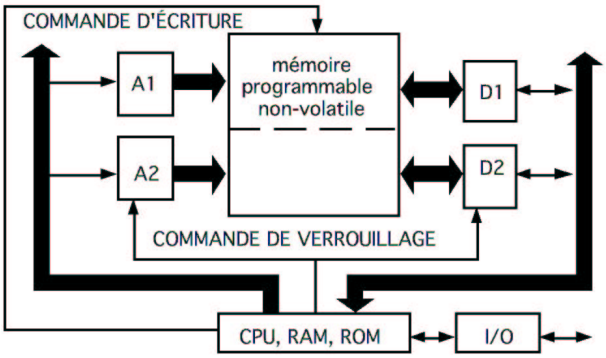

FIGURE 4 – Architecture du composant SPOM

tion de la mémoire, je décide donc de produire la toute première carte à une seule puce. La compagnie CII HB me procure fort aimablement des supports de contacts et des cartes en plastique. Je programme les algorithmes cryptographiques et les clés dans des circuits 8748 à quarante broches. Avec l´aide de la société Sorep à Chateaubourg, près de Rennes, une puce 8748 est extraite d'un circuit, puis connectée à un support en soudant les contacts utiles et, en novembre 1980, le microcircuit est collé avec succès dans une carte en plastique. Mais en raison du contexte politique (fin 80, début 81), la démonstration au plus haut niveau de l´État n´a pas eu lieu. La figure 3 montre également les restes de cette carte avec un peu de résine sur la puce 8748.

Le SPOM. Dans le microcalculateur auto programmable monolithique *(Self-Programmable One-chip Microcomputer, SPOM)*, le processeur contrôle lui-même tous les signaux électriques et logiques appliqués à la mémoire programmable non-volatile. Des registres s´interposent entre le bus général et les accès (adresse et données) à la mémoire programmable non-volatile. Ainsi les signaux imposés à la mémoire restent stables le temps nécessaire alors que le programme continue de se dérouler dans le processeur, ce qui entraîne une évolution des états sur le bus général. Illustrée par la figure 4, l'architecture du composant SPOM découverte par Ugon en 1977 donne lieu aux brevets « fondateurs » de la carte à processeur (voir la ramification dans l'arbre en figure 1).

La coopération entre Bull et Motorola se poursuit par la production du premier composant SPOM à partir d'avril 1981. Autour d'un cœur de 6805, il porte 36 octets en RAM, 1024 octets en EPROM et 1600 octets en ROM. Le premier masque sur ce composant est PC0, développé en collaboration avec le CCETT comme carte porte-clés d'accès conditionnel en radiodiffusion. Â partir de 1985, la société Eurotechnique (aujourd'hui ST Microelectronics) produit le second composant SPOM. Autour d'un cœur de 8049, il porte 44 octets de RAM, 1024 octets en EPROM et

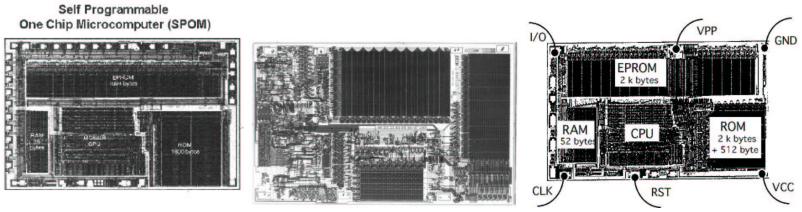

FIGURE 5 – Les trois premiers composants SPOM

2048 octets en ROM. Â partir de 1986, Motorola produit un troisième composant SPOM. Toujours autour d'un cœur de 6805, il porte 52 octets en RAM, 2048 octets en EPROM et 2048 octets en ROM (plus 512 octets en ROM pour des tests de fabrication). La figure 5 montre ces trois premiers composants SPOM.

2 Les débuts de la cryptologie dans le domaine public

L'informatique et la microélectronique ont d'abord été au service exclusif de la cryptanalyse (l'art de démanteler). La mise en œuvre de la cryptographie était alors une affaire de spécialistes : il fallait un « brevet de pilote » pour manipuler des machines mécaniques. Partageant tous les secrets de l'officier ou du diplomate, le chiffreur était un véritable « secrétaire ». L'apparition des machines à chiffrer électroniques altère irréversiblement la situation privilégiée du chiffreur en mettant l'informatique et la microélectronique au service de la cryptographie (l'art de mettre en œuvre).

2.1 L'algorithme DES et l'ouverture de la boîte de Pandore

Jusqu'aux années 70, quelques agences nationales se réservent la connaissance en cryptologie pour des usages militaire et diplomatique. Mais en 1974 et 1975, le bureau NBS *(National Bureau of Standards)* des États-Unis d'Amérique publie des appels à contributions pour un algorithme de chiffrement dans le journal officiel *(Federal Register)* afin de protéger les fichiers et les communications sensibles (mais non classifiées) des agences fédérales. En matière de cryptologie, la boîte de Pandore s'ouvre !

Avec l'agence NSA *(National Security Agency)* des États-Unis d'Amérique, la société IBM *(International Business Machine)* met au point l'algorithme DES *(Data Encryption Standard)* qui est publié en janvier 1977. Il permet de chiffrer et de déchiffrer un bloc de 64 bits sous le contrôle d'une clé secrète de 56 bits. Cette cryptographie est à *clé secrète*. On parle encore de *technique symétrique*, car la même clé permet de contrôler les deux opérations : le chiffrement et le déchiffrement.

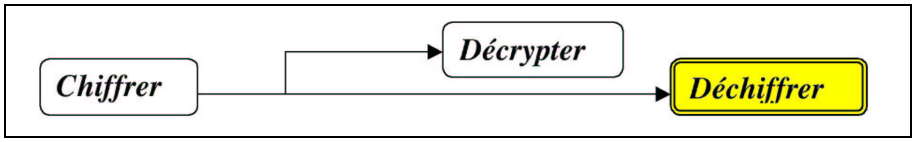

FIGURE 6 – Problématique de la confidentialité

2.2 Le concept de système à clé publique

Les consultations publiques sur l'algorithme DES intéressent beaucoup Hellman et Diffie, chercheurs à l'université de Stanford en Californie. En juin 1975 à Lennox au Massachusetts, puis, en juin 1976 à Ronneby en Suède, comme alternative à la cryptologie à clé secrète, ils suggèrent de nouvelles directions en cryptologie. En novembre 1976, ils publient leurs réflexions [4], introduisant ainsi le concept de « système à clé publique » pour assurer la confidentialité et/ou l'intégrité.

Note. Les deux chercheurs montrent l'intérêt des arithmétiques entières dans de grands ensembles finis. Beaucoup de chercheurs de par le monde, chacun sur l'ordinateur dont il dispose à l'époque, se constituent alors une bibliothèque arithmétique « multi précision ».
Note. Motivé par ma visite à Hellman à Stanford en avril 1977, au cours de laquelle Hellman me signale le nombre $r = 2^{121}.5^2.7^2.11^2.13.17.19.23.29.31.37.41.43.47.53.59$, j'écris de toute pièce une bibliothèque en Fortran sur l'ordinateur CII HB 10 070 de l'époque au CCETT. Je démontrer le 28 juin 1977 que les nombres $q' = r+1$ et $q = 2q'+1 = 2r+3$ sont premiers. Le nombre r s'écrit sur 198 bits et le nombre q sur 199 bits.
Note. Ma bibliothèque me permet de prouver la primalité de n'importe quel grand nombre p grâce au petit théorème de Fermat lorsque l'on connaît la décomposition complète de $p-1$ ou grâce aux suites de Lucas lorsque l'on connaît la décomposition complète de $p+1$ (voir Knuth [16]). Ainsi, le 14 septembre 1977, je démontre que $2^{210}-65$ et $2^{209}-33$ sont premiers en utilisant la décomposition complète de $2^{209}-32$.

La figure 6 illustre la problématique de la confidentialité. Une information peut atteindre quelqu'un d'autre que le légitime destinataire. Une communication peut être écoutée et interceptée. Le chiffrement consiste à remplacer le message clair par un cryptogramme et le déchiffrement à rétablir un message clair à partir d´un cryptogramme. Il convient de distinguer déchiffrement et décryptement. Celui qui décrypte cherche à rétablir le clair sans avoir connaissance a priori des secrets du déchiffreur. L'opération de décryptement n'admet pas d´inverse, de même que le verbe décrypter.

Lorsque la confidentialité est menacée dans un système transmettant de l'information dans le temps ou l'espace, au moins l'entité qui déchiffre doit protéger son opération en la gardant secrète. La figure 7 illustre la problématique de l'intégrité. Une information quelconque peut être injectée. Un message vraisemblable peut être forgé. Un message intercepté peut être modifié ou tout simplement retardé ou rejoué. L'émetteur légitime peut être simulé.

FIGURE 7 – Problématique de l'intégrité

Lorsque l'intégrité est menacée dans un système transmettant de l'information dans le temps ou l'espace, au moins l'entité qui prouve doit protéger son opération en la gardant secrète.

Par définition, j'appelle « contrôleur » l'ordinateur qui contrôle et « témoin » la structure, logicielle ou matérielle, qui détient, protège et utilise le secret au sein de l'ordinateur qui prouve.

— Pour identifier un utilisateur, le contrôleur doit le reconnaître. L'identification s'appuie sur le mot de passe et/ou la biométrie ; elle n'est pas cryptographique, même si la cryptographie renforce la sécurité de toute implémentation (par exemple, une suite d'images d'un mot de passe, voir Lamport [17]).
— Pour authentifier un ordinateur, le contrôleur doit être convaincu que l'ordinateur qui prouve (un serveur, un ordinateur personnel, une carte à puce) est autorisé à exécuter une action spécifique ou à accéder à des ressources, c'est-à-dire, qu'il représente véritablement un utilisateur autorisé.

Tout système à clé publique implique une clé privée et une clé publique. On peut facilement créer au hasard une paire de clés, mais on ne sait pas reconstituer la clé privée à partir de la clé publique.
– Pour assurer la confidentialité, la clé publique sert à chiffrer et la clé privée à déchiffrer.
– Pour assurer l'intégrité, la clé privée sert à prouver et la clé publique à contrôler.

On parle encore de *technique asymétrique*. Avec Davio et Quisquater, nous faisons le point [11] sur les techniques asymétriques en 1989. Suivant le schéma, un système à clé publique offre divers services : gestion des clés, chiffrement, authentification, signature, etc.

2.3 Le système RSA

En 1977, Rivest, Shamir et Adleman inventent le premier système à clé publique. En août 1977 dans la rubrique des jeux mathématiques sous la plume de Martin Gardner [8], la revue *Scientific American* publie le système RSA. En décembre 1977, le MIT *(Massachusetts Institute of Technology)* dépose une demande de brevet pour le système RSA aux États-Unis d'Amérique.

Note. Au vu de l'histoire du mémoire technique TM-82 racontée en 2.4 et de l'histoire du chiffrement sans secret racontée en 2.8, je m'interroge sur ce qui serait advenu si le MIT avait déposé une demande de brevet avant toute publication. La demande de brevet aurait probablement été classifiée et le système RSA porterait aujourd'hui un autre nom, après avoir été une fois de plus découvert ailleurs, sans doute en Europe.

Le système RSA repose sur la grande dissymétrie entre la construction du produit de grands facteurs premiers pris au hasard et la décomposition en facteurs premiers de grands nombres entiers composites. Chaque instance du système RSA s'appuie sur le problème de la factorisation d'un module n dont les facteurs premiers doivent rester secrets. Chaque couple de clés RSA comprend une clé privée $\langle s, n\rangle$ et une clé publique $\langle v, n\rangle$. Chaque clé RSA comprend un exposant, public ou privé, et un module public. L´opération élémentaire consiste à élever un argument de 0 à $n-1$ à la puissance indiquée, v ou s, dans l´anneau des entiers modulo n. Les deux opérations élémentaires sont inverses l´une de l´autre ; elles permutent l´ensemble des éléments de l´anneau. La simplicité de ces mécanismes élémentaires est à la base leur succès : le système RSA est si facile à expliquer et à assimiler.

À titre d'exemple dans l´article [8], la phrase « IT IS ALL GREEK TO ME » est transformée en un nombre entier selon un codage sommaire sans accent ni minuscule (A = 01, B = 02, ... Z = 26, l´espace étant codé par 00).

0920000919000112120007180505110019150013 05

Pour obtenir un cryptogramme, ce nombre est ensuite élevé à une certaine puissance, 9007 exactement, modulo un nombre appelé plus tard RSA-129 (car écrit sur 129 chiffres décimaux).

Note. Grâce à ma bibliothèque en Fortran sur l'ordinateur CII HB 10 070 de l'époque au CCETT, je vérifie l´exemple numérique dans la semaine suivant la publication de l'article au mois d´août 1977.

2.4 Le mémoire technique MIT/LCS/TM-82

L´article [8] propose la communication d´un mémoire technique, le fameux mémoire MIT/LCS/TM-82, avec des détails pratiques et des programmes informatiques en vue d´une éventuelle réalisation. Les inventeurs reçoivent plus de 3000 lettres, y compris une demande trois fois formulée par le CCETT : en septembre, en novembre et avec des vœux en décembre 1977.

Les 3000 demandes restent et resteront définitivement sans réponse. En effet, émue par les éventuelles conséquences d´une dispersion de la connaissance, la NSA interdit aux inventeurs de répondre et le mémoire technique TM-82 (daté du 4 avril 1977 selon Ellis, voir 2.8) est retiré des étagères du MIT dès la fin août 1977. J'apprend cela à Stanford et au MIT en avril 1978. En mars 1978, dans les *Communications of the ACM* [23], Rivest, Shamir et Adleman publient une version « simplifiée »

du mémoire technique TM-82, ce qui ouvre la voie à une longue controverse entre les libertés universitaires et le contrôle des publications en cryptologie.

2.5 Les vœux cryptographiques du CCETT à Rivest

En décembre 1977, afin d'attirer l'attention de Rivest sur la demande du TM-82 déjà formulée par le CCETT en septembre et en novembre, je décide d'adresser des vœux cryptographiques à Rivest.

HAPPY NEW YEAR BONNE ANNEE BLOAVEZ MAD LOUIS GUILLOU

Selon le codage sommaire de l'exemple, le nombre suivant représente le message en clair.

0801 16162500 14052300 25050118 00021514 14050001
14140505 00021215 01220526 00130104 00000012 15210919 00072109 12121521

Grâce à ma bibliothèque en Fortran, j'effectue les trois opérations suivantes.

1. J'élabore un couple de clés RSA pour la circonstance, similaire au couple de clés RSA du MIT. La fonction publique du CCETT est définie par l´exposant 10103 et le module public CCETT-129.
   114331674 69217002 15435665 48973781 01744179 08372627 88891346 00625638
   01820107 83582559 36575161 34887282 26552502 13131126 02207861 72216269
2. Le message en clair est d´abord signé par la clé privée du CCETT, ce qui donne le nombre suivant.
   072460412 34803883 80583418 12565585 28241025 03538873 48666078 03059292
   75891145 25792783 87392435 49891539 99792980 73981633 79601037 44428702
3. Le résultat est chiffré par la clé publique du MIT, ce qui donne le nombre suivant que je transmets.
   000866848 79834115 03924085 04327835 82608413 60733671 18725975 76718015
   02482627 37158155 63640755 92733287 50205408 75254272 47369834 22506663

Lors de ma visite au MIT en avril 1978, Rivest m'avoue n´avoir pas osé tester les vœux du CCETT car les ordinateurs disponibles à l'époque au MIT sont incapables de garantir un secret face à la curiosité des étudiants. En effet, pour tester ces vœux, il aurait fallu utiliser l´exposant privé du MIT, un secret permettant d'élucider un cryptogramme présenté dans l´article [8] en défi à la communauté scientifique, avec une promesse de 100 dollars au premier qui en rétablirait le clair.

2.6 La factorisation du nombre RSA-129

Selon le titre de l´article [8], il aurait fallu des millions d´années pour décomposer en facteurs premiers le nombre RSA-129. Pourtant, cela arrive seulement dix sept ans plus tard 8.

Sujet : RSA-129 Date : 27 avril 1994 04 :06 :25 GMT
Nous sommes heureux d´annoncer que
RSA-129 = 114381625757888867669235779976146612010218296721242362562561842 9
3570693524573389783059712356395870505898907514759929002687954354 1
*= 3490529510847650949147849619903898133417764638493387843990820577 **
32769132993266709549961988190834461413177642967992942539798288533
Le message chiffré publié est
9686961375462206147714092225435588290575999112457431987469512093081 62
9822514570835693147662288398962801339199055182994515781515 4
Ce message provient du chiffrement d´un message secret en utilisant l´exposant public 9007.
L´exposant secret est
1066986143685780244428687713289201547807099066339378628012262244966 31
0631259117744708733401685974623065539685445132771090536060 95
En déchiffrant le message chiffré avec l'exposant secret, on obtient
2008050013010709030023151804190001180500191721050113091908001519190906180107 05
Avec le code 01=A, 02=B, ..., 26=Z, et 00 pour l´espace, le message déchiffré se lit
THE MAGIC WORDS ARE SQUEAMISH OSSIFRAGE
Pour décomposer RSA-129 en facteurs premiers, nous avons utilisé la double variation de la méthode de factorisation par filtrage avec plusieurs polynômes quadratiques. Le filtrage a demandé environ 5000 Mips année ; il fut mené à bien en huit mois par 600 volontaires de plus de vingt pays, sur tous les continents, sauf l´antarctique.
Les relations partielles formaient une matrice presque vide de 569 466 rangées et 534 338 colonnes. Cette matrice fut réduite en une matrice dense de 188 614 rangées et 188 160 colonnes par élimination gaussienne structurée. Une élimination gaussienne ordinaire sur la matrice dense de 35 489 610 240 bits (4,13 Go) demanda 45 heures sur un ordinateur massivement parallèle 16K MasPar MP-1. Après trois premières dépendances infructueuses, la quatrième dépendance donna la décomposition de RSA-129.
Nous tenons à remercier tous ceux qui ont contribué par leur temps et leurs efforts à ce projet. Sans leur aide, cela n´aurait pas été possible.
Derek Atkins Michael Graff Arjen Lenstra Paul Leyland

FIGURE 8 – Annonce de la décomposition en facteurs premiers du nombre RSA-129

Rivest ayant remis un chèque de cent dollars à Lenstra, chacun des 600 contributeurs volontaires reçoit une photocopie du chèque aujourd'hui exposé, encadré sous verre, dans le bureau de Lenstra.

Ces vingt cinq dernières années, la décomposition des grands nombres en facteurs premiers a fait des progrès. Toutefois, ces progrès sont dus plus à l'évolution des matériels informatiques qu'à celle des algorithmes mathématiques. Les méthodes de 2002 avec les ordinateurs de 1977 sont bien moins performantes que les méthodes de 1977 avec les ordinateurs de 2002. Par conséquent, la confiance dont bénéficie le problème de la factorisation s'avère tout à fait méritée. Les tailles

conseillées pour les modules [18] sont 768 bits à court terme, 1024 bits à moyen terme et 2048 bits à long terme.

2.7 La consultation américaine

En 1982, le CCETT répond conjointement avec la compagnie CII HB à une consultation émise par le bureau NBS *(National Bureau of Standards)* des États-Unis d´Amérique sur les systèmes à clé publique. Notre proposition consiste à utiliser le système RSA sur des cartes à puce. Ce travail m'oblige à formaliser les programmes développés auparavant pour produire des couples de clés RSA et pour faire les diverses opérations nécessaires : chiffrer et déchiffrer, ou bien, prouver et contrôler ; il donne lieu pour moi à un brevet en commun avec Ugon (voir l'arbre de la figure 1). Â mon avis, la manœuvre américaine vise simplement à glaner de l'information à moindre frais et elle réussit : le bureau NBS ne donne aucune suite à la réponse.

2.8 La révélation différée du chiffrement sans secret (Non Secret Encryption)

Le 18 décembre 1997 lors d'une conférence à Cirencester, Cocks révèle que le chiffrement sans secret est secrètement à l'étude depuis 1969 au sein du groupe CESG *(Communications Electronics Security Group)* au GCHQ *(Government Communications Headquarters)* à Cheltenham, Gloucestershire.

Les papiers de l'époque sont alors déclassifiés. Depuis lors, ils sont disponibles sur le site du CESG.

— Dans un papier terminé en janvier 1970, à partir de considérations sur le bruit, Ellis établit la possibilité de « chiffrer sans secret ». Il découvre ainsi ce qui allait devenir le concept de clé publique.
— Le 20 novembre 1973, pour chiffrer sans secret, Cocks propose la fonction « puissance n-ième modulo n » : si n est le produit de deux grands facteurs premiers p_1 et p_2 que p_1 ne divise pas $p_2 - 1$ et p_2 ne divise pas $p_1 - 1$, alors la fonction permute l'anneau ; la permutation est inversée par la fonction « puissance x-ième modulo n » où x est le plus petit nombre entier tel que $\lambda(n) = \text{ppcm}(p_1 - 1, p_2 - 1)$ divise $x.n - 1$. Il découvre ainsi une instance particulière de ce qui allait devenir l'algorithme RSA.
— Le 21 janvier 1974, toujours pour chiffrer sans secret, Williamson propose un système « à double cadenas » utilisant la commutativité de l'exponentielle dans un corps fini.
— Le 10 août 1976, Williamson propose un système de mise à la clé utilisant les exponentielles modulaires dans de grands ensembles finis. Il découvre ainsi ce qui allait devenir l'algorithme DH.

— Un papier de Ellis en date de 1987 résume l'histoire du « chiffrement sans secret ». Ce papier confirme l'existence du mémoire technique MIT/LCS/TM-82 (voir 2.4) et le date du 4 avril 1977.

Note. Au vu de ces papiers, le chiffrement sans secret n'aborde que la problématique de la confidentialité (voir figure 6). Il semble que le groupe CESG n'ait pas entrevu la problématique de l'intégrité (voir figure 7). Pourtant sans signature numérique, il n'y a pas d'infrastructure de clés publiques.

Les travaux du groupe CESG ne donnent pas lieu à brevets pour au moins deux raisons.
– D'une part, au début des années 70, les algorithmes ne sont pas brevetables.
– D'autre part, le brevet est étranger, voire incompatible, avec les pratiques du GCHQ.

Entre temps, avec l'algorithme DES, la société IBM ouvre la voie aux brevets sur les algorithmes. En 1976, l'université de Stanford dépose une demande pour le concept de chiffre à clé publique et une demande pour l'algorithme DH. En 1977, le MIT dépose une demande pour l'algorithme RSA.

2.9 Les techniques sans transfert de connaissance (Zero Knowledge)

De façon générale, les techniques *Zero Knowledge* (sans transfert de connaissance, ZK) permettent d'une part, de prouver que l'on connaît une solution d'un problème complexe sans la révéler et d'autre part, de contrôler une solution d'un problème complexe sans en prendre connaissance. Mais le concept ZK est bien plus complexe à expliquer et à assimiler que le concept de clé publique.

Note. Les familles Guillou et Quisquater [21] se sont mises ensemble pour vulgariser le concept ZK : si l'on parvient à expliquer aux enfants, il est probable que le chef, voire le directeur, comprendra aussi.

Tout comme le RSA, les schémas suivants s'appuient sur le problème de la factorisation du module. Mais ils s'appuient aussi sur le concept ZK formalisé par Goldwasser, Micali et Rackoff [9], après une communication à Eurocrypt'84 dans laquelle Fisher, Micali et Rackoff [7] proposent un protocole où un nombre public est le carré modulaire d'un nombre privé pris au hasard. D'abord Fiat et Shamir [5, 6] proposent un schéma où chaque nombre privé est une signature Rabin-Williams [22, 25] de données d'identification ; le nombre de paires de nombres et le nombre d'itérations du schéma fixent ensemble un niveau de sécurité, par exemple, 2^{16} pour quatre paires et quatre itérations. Micali et Shamir [19] proposent un schéma où chaque nombre privé est la racine carrée modulaire d'un petit nombre premier. Avec Quisquater, nous [12, 13] proposons un schéma où le nombre privé est une racine modulaire v-ième, par exemple, une signature RSA

d'une identité ; le niveau de sécurité est $1/v$ avec une seule paire et une seule itération, par exemple, 2^{16} pour $v = 2^{16} + 1$. Ong et Schnorr [20], puis, Shoup [24], proposent un schéma où le nombre public est la puissance modulaire 2^k-ième d'un nombre privé pris au hasard.

À chaque exécution de chacun de ces schémas, le témoin produit d'abord un engagement (nombre positif inférieur au module) à partir d'un nouvel aléa selon une formule d'engagement, puis une réponse (nombre positif inférieur au module) à n'importe quel défi en utilisant un ou plusieurs nombres privés selon une formule de réponse. L'engagement, le défi et la réponse forment ensemble un triplet ZK. Par définition, un triplet ZK est valide quand il satisfait une formule de contrôle.

— En utilisant le ou les nombres privés selon le schéma, le témoin réussit chaque itération du schéma.
— Lorsque tous les défis possibles sont équiprobables, le contrôleur a exactement une chance sur le nombre total de défis de détecter une entité qui ne connaît aucun nombre privé. En effet, n'importe qui a une chance sur le nombre total de défis de deviner le défi ; il peut prendre une réponse au hasard et calculer un engagement grâce à la formule de contrôle. Mais toute entité anticipant un autre défi après s'être engagé connaît les réponses à deux défis pour le même engagement ; par définition, nous parlons d'une « paire entrelacée de triplets ZK ». Dès lors que la majorité des paires entrelacées de triplets ZK révèle le nombre privé ou un produit modulaire de nombres privés, une telle éventualité est incompatible avec le fait de ne connaître aucun nombre privé.
— Le contrôleur n'apprend rien sur la valeur du ou des nombres privés si ce n'est que le témoin les utilise. Il y a bien deux modes de production de triplets ZK : un mode privé et un mode public. En mode privé, la chronologie est d'abord l'engagement, puis la réponse à n'importe quel défi. En mode public, la chronologie est d'abord le défi et la réponse, puis l'engagement. Etant donné un triplet ZK produit au hasard, il n'y a pas moyen d'en déterminer le mode de production. Il n'y a pas de différence entre des données échangées durant l'exécution d'itérations du schéma et des données échangées entre deux entités ayant convenu à l'avance d'une liste de défis.

Note. Pour annuler le transfert de connaissance en authentification avec le système RSA, Brandt, Damgård, Landrock et Pedersen [2] préconisent de tirer un aléa de taille appropriée, de le hacher et d'appliquer la fonction RSA publique à la concaténation des deux champs pour obtenir un défi. La clé RSA privée sert à déchiffrer le défi, et si le code de hachage est correct, la réponse est l'aléa.

3 Les premiers développements mettant en œuvre des cartes à processeur

3.1 L'accès conditionnel en radiodiffusion

Publié au *Journal Officiel* du 20 mars 1978, le décret n^{o} 78379 aménage le monopole de radiodiffusion en France ; la diffusion de programmes spécifiques vers des publics limités, déterminés et identifiables devient possible par dérogation ; la voie de la radiodiffusion à péage est ouverte.

En juillet 1978, à la demande de la Direction Générale de Télédiffusion de France (TDF), un groupe d´action sur le contrôle de l´accès aux services diffusés est créé au sein du CCETT. En 1979, le groupe se transforme en un laboratoire « Cryptologie et Accès aux Services » (CAS).

J'assure dès le début l'animation du groupe, puis, du laboratoire, avec pour mission de concevoir, breveter et faire développer tous les éléments nécessaires pour assurer un péage à base de carte à puce.

En 1978, le CCETT conçoit un système complet pour contrôler l'accès aux magazines diffusés par le télétexte Antiope [10]. Le 6 février 1979, trois brevets sont déposés à mon nom. Ce sont des brevets « fondateurs » en matière d´accès conditionnel aux services audiovisuels : la plupart des systèmes de télévision à péage en appliquent aujourd'hui les principes.

— La carte d'accès de l'utilisateur est une carte porte-clés matérialisant des titres d'accès, c'est-à-dire, des clé et des droits d'accès associés. Le premier masque de carte porte-clés est appelé PC0.
— Les informations utiles de chaque magazine Antiope sont « embrouillées » et « désembrouillées » page par page, par « ou-exclusif » avec une suite d'octets chiffrants, produite par un générateur initialisé par un marquant (pour Antiope, le numéro de page) et une clé éphémère (rétablie par la carte d'accès).

Note. Un autre laboratoire de TDF à Issy les Moulineaux utilise une technologie éprouvée et disponible : la carte à pistes magnétiques, avec des secrets enterrés dans des décodeurs immatriculés.

3.2 La famille des cartes porte-clés d'accès conditionnel

L´accès conditionnel aux services audiovisuels présente quelques spécificités ne se retrouvant ni dans l´environnement bancaire, ni dans l´environnement téléphonique. En matière de cryptologie et de carte à puce, l´accès conditionnel a fait et fait encore figure de précurseur ; c´est un moteur de développements et une source d´innovations. Le piratage y est particulièrement virulent.

Des cartes porte-clés PC0 exécutent une paire d'algorithmes appelés « TDF » *(Twisted Double Field)* tout en mémorisant des clés d'exploitation C de 127 bits et des droits d'accès.

— Dans les cartes mères PC0, un algorithme calcule un cryptogramme E (127 bits) à partir d'un aléa X (66 bits), d'un paramètre P (24 bits), d'un message M (61 bits) et d'une clé C (127 bits).
— Dans les cartes filles PC0, un algorithme codé en 300 octets calcule un résultat R (61 bits) à partir d'un cryptogramme E (127 bits), d'un paramètre P (24 bits) et d'une clé C (127 bits) adressée par son nom (3 octets). Etant donnés un paramètre P et une clé C, le résultat R et le message M sont identiques.

Dans la carte fille PC0, une hiérarchie apparaît entre la clé unique de la carte et jusqu'à vingt clés d'opération avec des droits associés. Obtenue par diversification d'une clé maîtresse, la clé unique permet à l'émetteur de gérer des clés d'opération et des droits associés dans la carte fille. Les droits restreignent l'usage de la clé, par exemple, une période d'abonnement ou un crédit pour usage impulsif. Le paramètre est un critère d'accès, par exemple, la date à confronter à une période d'abonnement. Depuis bientôt quinze ans, le péage mis au point par le CCETT est exploité (aujourd'hui par Viaccess, filiale de France Telecom) avec la carte PC2 qui continue la famille des cartes « porte-clés ». Les cartes filles PC2 ne peuvent pas interagir entre elles ; elles ne réagissent que sous le contrôle de cartes mères PC2 spécialisées pour la gestion ou pour l'exploitation. Le système met aussi en œuvre des cartes mères d'émission pour émettre les cartes, et des cartes grands-mères pour gérer les clés dans les cartes mères. Pour chaque carte fille PC2, une seule clé apparaît en clair une seule fois dans le monde extérieur ; lors d'une écriture en fin de phase de fabrication des puces, c'est une clé temporaire de gestion, propre à la puce.

Note. C'est seulement en 1986 que l'algorithme DES apparaît dans un masque, à savoir le masque D1 de PHILIPS. Les ressources limitées des composants SPOM imposent une pression si forte aux programmeurs que sous la direction de Quisquater, ils réussissent la prouesse de programmer le DES sur environ 700 octets.

4 Les cartes bancaires à puce

De 1981 à 1983, le CCETT contribue largement à l'établissement des spécifications définitives de la carte bancaire française à puce, appelées B0 et publiées en janvier 1984.

De 1982 à 1984, les banques françaises regroupées au sein du « GIE Carte à Mémoire » mènent trois expérimentations dans trois villes : Blois, Caen et Lyon. L´objectif est de tester la viabilité technique et économique de la carte à puce

dans la vie réelle avant d´établir des spécifications définitives. Appelée IPSO, l'expérimentation portait sur 750 terminaux et 125 000 cartes. La société Flonic-Schlumberger distribue une carte à logique câblée à Lyon. La société Philips qui a une usine à Caen y distribue des cartes à deux puces. La compagnie CII HB distribue ses cartes à une seule puce à Blois, avec le composant SPOM produit par Motorola.

Â l´issue des trois expériences, seule l´expérience de Blois est jugée convaincante et les banques choisissent la carte à processeur de Bull plutôt que la carte à logique câblée de Schlumberger. En 1985, une commande de 16 millions de cartes initialise la généralisation des puces dans les cartes bancaires françaises.

En 1994, un montant de 807 milliards de Francs est traité au cours de 2,35 milliards de transactions exécutées par 22,8 millions de cartes. La fraude se chiffre à 0,035 %, répartie entre 0,032 % en France (technologie puce) et 0,20 % à l'étranger (technologies embossage et piste magnétique). En 1994, la carte à puce est un succès français.

5 L'authentification statique des cartes à puce

Le CCETT conçoit l'authentification « statique » des cartes à puce en 1983. Les spécifications de la carte bancaire à puce de janvier 1984 comprennent trois nombres produits en août 1983 par une machine « magnolias » développée au laboratoire CAS du CCETT. Ces trois nombres de 321 bits sont prévus pour une durée de cinq à dix ans, durée qui n'a pas été respectée.

— Un nombre opérationnel : les facteurs sont disponibles sur l'Internet depuis l'automne 1998, suite à la décomposition en facteurs premiers par Humpich. L'affaire fait grand bruit.
— Un nombre de secours : les facteurs sont égarés au cours de déménagements successifs dans Paris, la dissolution du groupement de la carte à mémoire et la création du groupement des cartes bancaires.
— Un nombre de test : les facteurs font partie des spécifications.

Note. Durant l'été 2000 au laboratoire d'informatique de Polytechnique, Morain à ma demande décompose les trois nombres de 321 bits. En conclusion, un ordinateur personnel cadencé à 400 MHz factorise un nombre de 321 bits en une semaine.

Durant quelques années à partir de 1984, un dispositif « camélias » développé au laboratoire CAS du CCETT calcule une valeur d'authentification lors de la personnalisation de chaque carte : la valeur d'authentification est une signature RSA de données d'identification de la carte.

Comme les mêmes données sont échangées à chaque utilisation de la carte, le dialogue est statique. Si elle renforce une authentification visuelle en local dans la bou-

tique du marchand, une telle méthode est inutile sur un automate sans présence humaine, ou bien à distance, par exemple, à travers l'Internet.

Note. L'authentification statique fait encore partie des spécifications internationales publiées par Europay, Mastercard et Visa pour la première fois en 1996 sous le nom EMV'96 et mises régulièrement à jour depuis. Nous ne considérerons pas plus avant cette méthode statique où la carte ne met aucun secret en œuvre.

5.1 L'authentification dynamique des cartes à puce

Les spécifications EMV imposent le RSA pour l'authentification dynamique. Malheureusement, le nombre d'opérations requis pour exécuter la fonction RSA privée impose de dédier une partie de la puce à un opérateur arithmétique d'exponentiation modulaire, opérateur appelé « crypto processeur ».

Le schéma GQ1 [12, 13] est aujourd'hui utilisé à très grande échelle ; il est à la base du produit NetWare de Novell : cent millions d'utilisateurs (un secret à long terme est utilisé pour calculer un secret à court terme durant la phase d'entrée en session). Le schéma GQ1 est une preuve ZK de connaissance d'une signature RSA. Au moment de la découverte, avec Quisquater, nous étions convaincus de détenir « la » méthode d'authentification dynamique des cartes ; l'histoire nous prouva que non.

Avec Quisquater, nous proposons aujourd'hui le schéma GQ2 ; c'est une preuve ZK de connaissance de décomposition du module n. GQ2 est beaucoup plus efficace que RSA : par exemple, pour un module de 1024 bits, le ratio est de 30 à 40 en authentification pour un défi de 16 bits et de 10 à 13 en signature pour un défi de 80 bits. Comme GQ2 admet une preuve de sécurité, la réduction de calcul n'affecte pas la sécurité. Une première version parait en 2001 [15]. Aujourd'hui, à partir de n'importe quelle clé RSA privée, on sait établir une paire de clés GQ2. Pour utiliser un module RSA en GQ2, il n'y a rien à ajouter aux certificats existants car il n'y a pas à certifier les ajouts GQ2 publics.

Note. Soient deux grands facteurs premiers dont le produit forme le module n. Chaque facteur premier p détermine un nombre a tel que $p-1$ est divisible par 2^a mais pas par 2^{a+1} ; le plus grand des nombres aa détermine le paramètre d'adaptation b ; un paramètre de sécurité k et un paramètre de multiplicité m fixent ensemble un niveau de sécurité, par exemple, 2^{16} pour $m = 2$ et $k = 8$. La clé GQ2 publique comprend le module n et les ajouts suivants : les paramètres b, k et m, ainsi que m nombres de base figurant parmi les 54 premiers nombres premiers (le 55-ième nombre premier est 257). Outre le module n et les paramètres b, k et m, la clé privée GQ2 comprend m nombres privés ; chacun est une racine modulaire $2^{(}k+b)$-ième du b-ième carré d'un nombre de base.

Note. Pour qu'une instance GQ2 soit équivalente avec la factorisation du module n, il suffit qu'au moins un nombre de base soit tel que les symboles de Legendre par rapport à p_1 et p_2 soient différents si $b_1 = b_2$ ou que le symbole de Legendre par rapport à b_i soit -1 si $b_i > b_{3-i}$.

RSA statique	**RSA dynamique**
Voici mon identité *Id* (l'autorité m'a fourni une signature RSA de *Id*)	Voici un certificat (signé par l'autorité) liant mon identité *Id* à une clé publique RSA
Je prouve en révélant la signature RSA de *Id*	Je prouve en inversant ma permutation RSA pour déchiffrer n'importe quel défi
GQ1	*GQ2*
Voici mon identité *Id* (l'autorité m'a fourni une signature RSA de *Id*)	Voici un certificat (signé par l'autorité) liant mon identité *Id* à un module n
Je prouve que je connais la signature RSA de *Id* sans en révéler les facteurs (par ZK)	Je prouve que je connais la décomposition de n sans en révéler la valeur (par ZK)

FIGURE 9 – Synthèse des schémas RSA et GQ

Note. Si le module n est congru à 1 modulo 4 avec $b = 1$ et si le symbole de Jacobi d'un nombre de base par rapport à n est -1, alors l'équivalence avec la factorisation est évidente.

La figure 9 résume les schémas RSA statique et dynamique, ainsi que GQ1 et GQ2. Dans les quatre cas, l'autorité utilise sa clé RSA privée et le contrôleur utilise la clé RSA publique correspondante.

6 La normalisation

La norme complète le brevet : le brevet exige d'expliquer tout le savoir-faire pour le publier ; la norme exige d'octroyer licence du brevet à tout concurrent à des conditions raisonnables et non discriminatoires, *under fair and reasonable conditions* selon l'expression consacrée. La normalisation joue un rôle essentiel pour stabiliser les technologies, assurer la concurrence entre les constructeurs, et donc en fin de compte, permettre l'exportation des produits et des services.

6.1 L'accès conditionnel en radiodiffusion

De 1983 à 1987 en matière de télévision à péage, l'UER (Union Européenne de Radiodiffusion) établit des spécifications d'accès conditionnel avec un vocabulaire approprié (voir figure 10).

Le mot de contrôle (CW, *Control Word*) sert de clé d'embrouillage et de désembrouillage ; le message de contrôle de titre d'accès (ECM, *Entitlement Control Message*) comprend un nom, un paramètre et une paire de cryptogrammes, correspon-

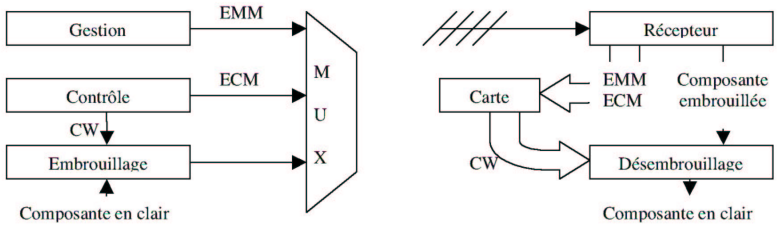

FIGURE 10 – Diagramme d'un système d'accès conditionnel

dant aux trois variables externes de l'algorithme TDF. Le titre d'accès *(Entitlement)* désigne une clé d'opération et les droits associés ; il est matérialisé par la carte. En télévision, le signal lui-même indique la parité de l'usage du mot de contrôle, avec un mot de contrôle « pair » et un mot de contrôle « impair ». Des messages de gestion de titre d'accès (EMM, *Entitlement Management Message*) permettent de gérer les titres dans les cartes des usagers.

Chaque message de contrôle (ECM) ou de gestion (EMM) comprend deux ou trois champs :

— un ou plusieurs éléments de données, dont un nom et des conditions d'accès ou des actions de gestion,
— zéro, un ou plusieurs cryptogrammes,
— et enfin un « code d'authentification du message » (MAC, 64 bits).

La carte rejette le message si le code MAC est incorrect. La carte poursuit le traitement si le code MAC est correct, généralement en exécutant l'une des actions suivantes :
– le calcul d'une paire de mots de contrôle (le courant et le suivant),
– l'écriture d'une clé (typiquement, une clé d'exploitation),
– la mise à jour de droits d'accès (période d'abonnement ou crédit).

Aujourd'hui, tous les opérateurs de télévision par satellite utilisent les normes EP-DVB établies de 1992 à 1996 dans le cadre du projet européen DVB *(Digital Video Broadcasting)*. Elles comportent le codage MPEG2 *(Mobile Picture Expert Group)* pour l'image, et un système d'accès conditionnel avec l'algorithme commun d'embrouillage *(Common Scrambling Algorithm)* pour le péage.

Le projet EP-DVB a repris tout le vocabulaire développé par l'UER de 1983 à 1987 à partir de contributions du CCETT.

6.2 Les techniques de sécurité

En 1981, l'Organisation Internationale de Normalisation (ISO) décide de normaliser des techniques cryptographiques au sein du TC97 / WG1 jusqu'en 1989. Ensuite, l'ISO/CEI poursuit la normalisation de techniques cryptographiques au sein de l'ISO/CEI JTC1 / SC20 jusqu'en 1991, puis, la normalisation de techniques de sécurité au sein de l'ISO/CEI JTC1 / SC27. Ces structures successives illustrent les problèmes politiques posés par l'introduction de la cryptologie dans le domaine public.

Le document ISO / DIS 8227, *Information processing, Data encipherment, Specification of algorithm DEA1*, est mis en circulation pour un vote de six mois le 13 juin 1985. Tous les Membres P du SC20 votent *oui* à l'exception de l'ANSI qui vote *non*. Ensuite l'ANSI émet un veto au Conseil de l'ISO qui abandonne donc le projet. Ainsi, la première tentative de normaliser le DES échoue.

De 1989 à 1993, j'assure l'animation de deux groupes de travail successifs : SC20 / WG2, *Techniques à clé publique*, puis SC27 / WG2, *Mécanismes de sécurité*.

En 1991, le document DIS 9796, *Schéma de signature numérique rétablissant le message*, est accepté par tous les membres P du SC27, sauf l'ANSI. Les États-Unis suggèrent alors un veto au Conseil de l'ISO, comme en 1985 pour le DES. Â Tokyo en avril 1991, durant la fin de semaine entre les réunions des groupes de travail et la réunion du SC27, sous le contrôle de Jahl, Président du SC27 à l'époque, je prépare un document où les spécifications de l'algorithme RSA figurent en annexe informative. L'ANSI accepte ce compromis et l'ISO publie la norme ISO/CEI 9796 en 1991.

En 1999, la norme ISO/CEI 9796 :1991 est retirée suite à une attaque [3] contre le mécanisme de format. Censé éliminer les attaques dues aux propriétés multiplicatives du RSA, le mécanisme de format joue en effet un rôle essentiel dans la sécurité de toute mise en œuvre du schéma RSA en signature.

Aujourd'hui, j'assure l'édition de deux normes : ISO/CEI 9798-5, *Authentification d'entité, Mécanismes basés sur des techniques sans transfert de connaissance*, et ISO/CEI 14888-2, *Signature numérique, Mécanismes basés sur la factorisation de nombres entiers*.

Ces deux normes comprennent les schémas GQ1 et GQ2. Elles comprennent aussi le schéma RSA, avec, dans la norme de signature, un mécanisme de format appelé PSS *(Probabilistic Signature Scheme)* dû à Bellare et Rogaway [1] ; la sécurité du schéma de signature RSA-PSS est prouvée.

Dans la pratique, les schémas ZK (et les schémas de signature numérique) mettent en œuvre une fonction de hachage, telle que spécifiée dans la série de normes ISO/CEI 10118, *Fonctions de hachage*.

— Un premier usage élimine toute possibilité de détourner le témoin en oracle de signature. Avant de s'engager, le témoin doit alors avoir reçu un engagement de défi. Ce pas préliminaire du protocole assure l'indépendance entre le tirage de l'aléa par le témoin et le tirage du défi par le contrôleur.
— Un deuxième usage consiste à remplacer la transmission de l'engagement par la transmission d'un code de hachage faisant intervenir l'engagement et un message, de façon à authentifier non pas seulement une entité, mais aussi un message.
— Un troisième usage permet de lier un ou plusieurs triplets ZK à un message de façon à former une signature numérique. Le code de hachage du message et d'un ou plusieurs engagements donne un ou plusieurs défis. L'entropie totale de ces défis doit être suffisante, c'est-à-dire, 80 bits ou plus.

6.3 L'interface des cartes à puce

En octobre 1981, l'Organisation Internationale de Normalisation (ISO) décide également de normaliser l'interface des cartes à puce. Comme j'avais déjà dû spécifier cette interface dans le cadre des travaux sur l'accès conditionnel, je participe activement aux activités correspondantes dès le début, d'abord au sein de l'ISO TC97 / SC17, puis à partir de 1989, au sein de l'ISO/CEI JTC1 / SC17.

Bien souvent en normalisation, les nouveaux arrivants tentent de casser l'avance des défricheurs par des normes remettant en cause les développements existants. Trois tentatives de ce genre de déstabilisation émaillent l'histoire de la normalisation des cartes à puce.
– Une attaque a réussi : la position des contacts sur la carte.
– Une attaque a échoué : la fréquence de référence pour l'horloge de la carte.
– Une attaque a été désamorcée : les protocoles d'échange T = 0 et T = 1.

Position des contacts. Un accord unanime est rapidement obtenu sur le type (en surface et non pas sur le bord), sur la forme (une surface rectangulaire minimum), sur la position relative et sur les fonctions. Cet accord partiel pérennise la fabrication des puces et des microcircuits (montage de puces sur supports de contacts, voir figure 2). Il reste à positionner le microcircuit sur la carte, en tenant compte des normes d'estampage et de pistes magnétiques, ainsi que de particularités nationales. La position « haute », couramment utilisée en France, devient « transitoire » dans la norme avant de disparaître ; la position basse a donc gagné cette première bataille. L'attaque vise les fabricants de cartes, et non pas les fabricants de puces. La figure 11 récapitule les positions.

Fréquence de référence. La carte doit recevoir la fréquence de référence sur le contact d'horloge CLK pour que les données soient échangées à 9 600 bits/s sur le contact I/O. La société Motorola considère que 4 MHz est un seuil technologique

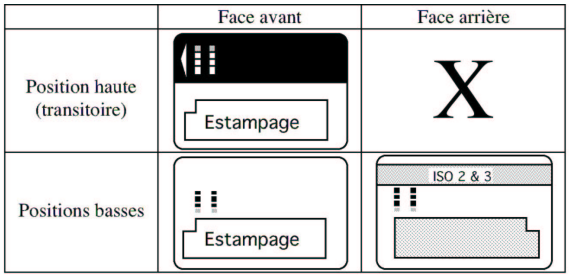

FIGURE 11 – Positions des contacts sur une carte à puce

dans le prix des puces. À environ 10 % en dessous de ce seuil, la fréquence la plus répandue dans le monde est 3,579545 MHz, utilisée par les téléviseurs NTSC. Un bit dure alors environ 372 coups d'horloge. Les Japonais et les Allemands critiquent cette fréquence de « télévision ». Ils suggèrent une fréquence « informatique », selon eux plus appropriée, à savoir 4,9152 MHz, car il faut alors exactement 512 coups d'horloge pour constituer un bit. Leur suggestion n'est pas été retenue. L'attaque vise cette fois les fabricants de puces.

Protocoles d'échange. Le protocole T = 0 fixe l'échange de caractères asynchrones sur le contact I/O en mode semi duplex. La détection d'erreur et la répétition sont gérées au niveau du caractère par un signal d'erreur à la fin de chaque caractère litigieux. Le protocole T = 1 fixe l'échange de blocs de caractères asynchrones également en mode semi duplex. La détection d'erreur et la répétition sont gérées au niveau de chaque bloc. Pour être indépendante du protocole de transmission, la commande se décline en APDU *(Application Protocol Data Unit)*, avec des règles d'insertion dans les TPDU *(Transmission Protocol Data Unit)* T = 0 et T = 1. Par conséquent, on peut aujourd'hui concevoir et fabriquer des coupleurs avec une interface APDU du côté de l'ordinateur et les protocoles T = 0 et T = 1 du côté de la carte. L'ordinateur n'a pas besoin de savoir quel est le protocole utilisé par la carte. Puisque les deux protocoles offrent alors le même service, il vaut mieux utiliser la carte en protocole T = 0 (environ 180 octets de programme et deux octets en mémoire tampon) plutôt que le protocole T = 1 (au moins 700 octets de programme et deux blocs en mémoire tampon). Cette attaque vise à dénigrer T = 0 et à promouvoir un protocole « pur », T = 1, marchant au pas de l'oie, bloc après bloc.

Dans le cadre de la normalisation mondiale de l'interface des cartes à puce, j'assure l'édition de deux normes en cours de révision : ISO/CEI 7816-3, *Cartes à circuit intégré, Cartes à contacts : Interface électrique et protocoles de transmission*, et ISO/CEI 7816-4, *Cartes à circuit intégré, Organisation, sécurité et commandes pour les échanges*. Une phase de réorganisation de la série ISO/CEI 7816 s'achève, avec rectification des frontières entre les parties 3 à 6, 8 et 9. Les parties 1, 2, 3, 10 et 12

sont spécifiques aux cartes à contacts. Les parties 4 à 9, 11, 13 et 15 s'appliquent à toute carte, aussi bien à contacts que sans. J'estime que les cartes à puce, et donc les normes, seront bien plus utilisées à l'avenir qu'elles ne l'ont été jusqu'à ce jour. Cette réorganisation est un acte de foi en l'avenir.

Mes travaux de normalisation en accès conditionnel et en techniques de sécurité ont sensiblement influencé mon action au niveau de l'interface des cartes. Voici deux exemples au niveau de la norme ISO/CEI 7816-4 : la messagerie de sécurité et la commande d'authentification générale.

Messagerie de sécurité. Pour assurer la protection des données échangées avec une carte, dès la première version de la norme ISO/CEI 7816-4 :1995, *Cartes à circuit intégré, Commandes intersectorielles pour les échanges*, j'ai introduit une messagerie de sécurité *(Secure Messaging)*. Cette messagerie de sécurité généralise la structure des messages ECM et EMM définie dans la norme UTE C 90-005 de novembre 1991.

Commande d'authentification générale. Les deux commandes d'authentification existantes, à savoir, « *authentification externe* » et « *authentification interne* », utilisent un protocole à deux échanges : un défi suivi d'une réponse. Pour remédier à cette limitation, j'ai introduit une nouvelle commande « *authentification générale* » dans le cadre de la révision de la norme ISO/CEI 7816-4, *Cartes à circuit intégré, Organisation, sécurité et commandes pour les échanges* ; elle permet d'accommoder la norme ISO/CEI 9798-5, en particulier les protocoles à trois échanges : engagement, défi et réponse.

Dans le cadre de la commande « *authentification générale* », chaque champ de données, en commande comme en réponse, doit comporter un objet de données construit (formulaire, *template*) pour regrouper des « *objets de données pour l'authentification dynamique* ». La figure 12 indique les objets de données pour l'authentification dynamique.

Une fonction d'authentification se traduit par une chaîne de commandes « *authentification générale* ». Le premier champ de données en commande indique la fonction comme suit.

— Un engagement vide ou un code d'authentification vide indique une « authentification interne ». Un engagement ou un code doit venir en réponse.
— Un défi vide indique une « authentification externe ». La commande doit également comporter un engagement. Un défi doit venir en réponse.
— L'absence d'objet vide indique une « authentification mutuelle ». La réponse doit comporter les mêmes objets que la commande.

Plus généralement, un objet vide exprime une demande : si un objet est vide dans le formulaire d'un champ de données, il doit être rempli dans le formulaire du prochain champ de données.

Etiquette	Valeur
'7C'	Ensemble d'objets de données avec les étiquettes suivantes
'80'	Engagement (un ou plusieurs nombres positifs et inférieurs au module)
'81'	Défi (un ou plusieurs nombres, positifs ou nuls, inférieurs à l'exposant public)
'82'	Réponse (un ou plusieurs nombres positifs et inférieurs au module)
'83'	Engagement de défi (code de hachage d'un aléa où figurent un ou plusieurs défis)
'84'	Code d'authentification obtenu par hachage à partir d'un ou plusieurs champs de données M et d'un objet d'engagement W : $h(M,W)$ ou $h(h(M),W)$ ou $h(h(M),h(W))$ ou $h(M,h(W))$
'85'	Exponentielle (nombre positif pour établir une clé de session par une technique d'agrément)
'A0'	Ensemble d'objets de données pour l'identification
Dans ce contexte, l'ISO réserve les autres objets de la classe « spécifique au contexte » (premier octet de '80' à 'BF')	

FIGURE 12 – Objets de données pour l'authentification dynamique

En cas d'authentification mutuelle, une paire d'objets « exponentielle » permet l'établissement d'une clé de session (à la « Diffie Hellman »).

Pour protéger des données échangées durant une session, chaque entité doit maintenir un code de hachage mis à jour à chaque champ de données « sensibles ». Le contrôleur reconstruit d'abord un engagement, puis, un code de hachage.

7 Conclusion

Si la carte à puce est aujourd'hui considérée comme une invention française, c´est simplement que l'arbre a pris racine (voir figure 1) en France ; les développements industriels y ont commencé à la suite d'un invraisemblable enchaînement d'évènements.

En 1992, l'Institut IEEE *(Institute of Electrical and Electronics Engineers)* publie un livre intitulé *Contemporary Cryptology, the Science of Information Integrity*. L'éditeur et co-auteur Simmons a tenu à ce que le livre comporte un chapitre sur la carte à puce et que ce chapitre soit rédigé en Europe sous ma responsabilité, avec Ugon et Quisquater [14]. Dans la préface, il écrit en prophétie : « Cette application (carte à puce) mettra un instrument sophistiqué, dédié à l'intégrité de l'information, dans la poche de pratiquement chaque personne dans le monde, et sera probablement l'application la plus répandue, jamais réalisée, de schémas cryptographiques. »

La production mondiale de composants SPOM progresse de 30 millions en 1994 à 375 millions en 1997. En 1997, Europay, Mastercard et Visa annoncent que « Toutes les cartes bancaires du monde seront à puce en l'an 2000 » et que « Les pistes magnétiques seront abandonnées en 2002 ». En 1997, la production de cartes à puce est estimée entre 1 et 2 milliards pour l'an 2000. La carte à puce serait alors un succès mondial après le succès français de 1994 et le succès européen de 1997.

L'avènement de la prophétie de Simmons s'avère un peu plus laborieux que prévu. La production est seulement de 541 millions pour l'an 2000, en progression de 36 % sur un an, se répartissant en 370 millions pour la téléphonie mobile (+85 %), 120 millions pour le secteur bancaire, 25 millions pour la télévision à péage, 20 millions pour la santé et 3 millions pour le transport, avec comme répartition géographique : 55 % pour l'Europe et seulement 4 % pour l'Amérique du Nord.

Il apparaît que la téléphonie mobile a sauvé la carte à puce qui a été au bord de l'asphyxie. Cependant, je suis convaincu que la puce va se généraliser dans les cartes bancaires et que la sécurité des ordinateurs et des réseaux privés virtuels passera par la carte à puce ou plutôt un « jeton cryptographique » portant et utilisant les droits d'accès de l'utilisateur, en particulier, tous les secrets pour s'authentifier, signer et gérer des clés. L'histoire de la carte à puce n'en est encore qu'à ses débuts.

On attend toujours une carte sans contact avec authentification dynamique, car la consommation de la puce pose alors un problème particulièrement contraignant. Je pense que la suite de l'histoire éclaircira le rôle du concept ZK dans l'authentification dynamique des cartes sans contacts. Je pense qu'il y a là un champ d'utilisation de GQ2.

Références bibliographiques

[1] M. Bellare, P. Rogaway, « The exact security of digital signatures : How to sign with RSA and Rabin », *Proc. Eurocrypt '96*, U. Maurer, Ed., Lecture Notes in Computer Science, Vol 1070, Advances in Cryptology, p. 399-416, Berlin, Springer-Verlag, 1996.

[2] J. Brandt, I. Damgård, P. Landrock, T. Pedersen, « Zero-knowledge authentication scheme with secret key exchange », *Proc. Crypto '88*, Sh. Goldwasser, Ed., Lecture Notes in Computer Science, Vol 403, Advances in Cryptology, p. 583-588, Berlin, Springer-Verlag, 1990.

[3] J.-S. Coron, D. Naccache, J.P. Stern, « On the security of RSA padding », *Proc. Crypto '99*, M. Wiener, Ed., Lecture Notes in Computer Science, Vol 1666, Advances in Cryptology, p. 1-12, Berlin, Springer-Verlag, 1999.

[4] W. Diffie, M. Hellman, « New directions in cryptography », *IEEE Transactions on Information Theory*, Vol IT-22, p. 644-654, novembre 1976.

[5] U. Feige, A. Fiat, A. Shamir, « Zero knowledge proofs of identity », *Journal of Cryptology*, Vol 1, p. 77-94, 1988.

[6] A. Fiat, A. Shamir, « How to prove yourself : Practical solutions to identification and signature problems », *Proc. Crypto '86*, A.M. Odlyzko, Ed., Lecture Notes in Computer Science, Vol 263, Advances in Cryptology, p. 186-194, Berlin, Springer-Verlag, 1987.

[7] M.J. Fisher, S. Micali, C. Rackoff, « A secure protocol for oblivious transfer », *Journal of Cryptology*, Vol 9-3, p. 191-195, 1996 (présenté à Eurocrypt '84, mais publié douze ans plus tard).

[8] M. Gardner, « A new kind of cipher that would take millions of years to break », *Scientific American*, Vol 237-8, p. 120-124, 1977.

[9] Sh. Goldwasser, S. Micali, C. Rackoff, « The knowledge complexity of interactive proof systems », *SIAM Journal on Computing*, Vol 18, p. 186-208, 1989.

[10] L.C. Guillou, « Des techniques de sécurité pour l'exploitation du réseau Antiope », *Actes du Troisième Congrès FBVI/FAIB*, Bruxelles, Belgique, 5-6 octobre 1982, p. 135-143.

[11] L.C. Guillou, M. Davio, J.-J. Quisquater, « Public-key techniques : randomness and redundancy », *Cryptologia*, Vol 13-2, p. 167-189, 1989.

[12] L.C. Guillou, J.-J. Quisquater, « A practical zero-knowledge protocol fitted to security microprocessor minimizing both transmission and memory », *Proc. Eurocrypt '88*, C.G. Günther, Ed., Lecture Notes in Computer Science, Vol 330, Advances in Cryptology, p. 123-128, Berlin, Springer-Verlag, 1988.

[13] L.C. Guillou, J.-J. Quisquater, « A paradoxical identity-based signature scheme resulting from zero-knowledge », *Proc. Crypto '88*, Sh. Goldwasser, Ed., Lecture Notes in Computer Science, Vol 403, Advances in Cryptology, p. 216-231, Berlin, Springer Verlag, 1988.

[14] L.C. Guillou, M. Ugon, J.-J. Quisquater, « The smart card, a standardized security device dedicated to public cryptography », *Contemporary Cryptology, the Science of Information Integrity*, Chap 13, p. 561-613, IEEE press, Piscataway, 1992.

[15] L.C. Guillou, M. Ugon, J.-J. Quisquater, « Cryptographic authentication protocols for smart cards », *Computer Networks Magazine*, Vol 36, p. 437-451, North Holland Elsevier Publishing, juillet 2001.

[16] D.E. Knuth, *he Art of Computer Programming*, Volume 2. Addison-Wesley, 3rd edition, 1997.

[17] L. Lamport, « Password identification with insecure communications », *Communications of the ACM*, Vol 24-11, p. 770-772, 1981.

[18] A.K. Lenstra, E.R. Verheul, « Selecting cryptographic key sizes », *Journal of Cryptology*, Vol 14, N 4, p. 255-293, 2001.

[19] S. Micali, A. Shamir, « An improvement of the Fiat-Shamir identification and signature scheme », *Proc. Crypto '88*, Sh. Goldwasser, Ed., Lecture Notes in Computer Science, Vol 403, Advances in Cryptology, p. 244-247, Berlin, Springer Verlag, 1988.

[20] H. Ong, C.P. Schnorr, « Fast signature generation with a Fiat-Shamir-like scheme », *Proc. Eurocrypt '90*, I.B. Damgård, Ed., Lecture Notes in Computer Science, Vol 473, Advances in Cryptology, p. 432-440, Berlin, Springer Verlag, 1991.

[21] J.-J., M., M. and M. Quisquater, L.C., M.-A., G., A., G. and S. Guillou, with the help of T. Berson, « How to explain zero-knowledge protocols to your children », *Proc. Crypto '89*, G. Brassard, Ed., Lecture Notes in Computer Science, Vol 435, Advances in Cryptology, p. 628-631, Berlin, Springer Verlag, 1990.

[22] M.O. Rabin, « Digital signatures and public-key functions as intractable as factorization », *Technical Report MIT/LCS/TR-212*, MIT Laboratory for Computer Science, January 1979.

[23] R.L. Rivest, A. Shamir, L. Adleman, « A method for obtaining digital signatures and public-key cryptosystems », *Communications of the ACM*, Vol 21-2, p. 120-126, 1978.

[24] V. Shoup, « On the security of a practical identification scheme », *Proc. Eurocrypt '96*, U. Maurer, Ed., Lecture Notes in Computer Science, Vol 1070, Advances in Cryptology, p. 344-353, Berlin, Springer Verlag, 1996.

[25] H.C. Williams, « Some public-key crypto-functions as intractable as factorization », *Proc. Crypto '84*, G.R. Blakley and D. Chaum, Eds., Lecture Notes in Computer Science, Vol 196, Advances in Cryptology, p. 66-70, Berlin, Springer Verlag, 1985.

Biographie de l'auteur

Louis Guillou est expert émérite à la Division R&D de France Telecom. En février 1973, il fut embauché au CCETT à Rennes par l'ORTF. Il développa la cryptologie dans les cartes à puce pour créer des systèmes de télévision à péage. Il a joué et joue encore un rôle important en normalisation de l'interface des cartes et des techniques cryptographiques à l'ISO/CEI, ainsi que des systèmes d'accès conditionnel

à l'UER/EBU, à l'UIT-T et à l'EP-DVB. Il a déposé plus de 25 demandes de brevets. Il est auteur ou co-auteur de plus de 50 publications. Il a servi comme président du programme pour Eurocrypt '95 à St Malo. Il est ingénieur de l'Ecole Centrale des Arts et Manufactures (ECP 70) et Docteur Ingénieur en Electronique appliquée aux Télécommunications (Rennes, 1973). Il est également très fier de son Diplôme Supérieur d'Etudes Celtiques (Rennes, 1974).