Histoire de la Cryptologie et problématiques actuelles

(présentation du 26 mars lors de l'AG d'ESTEL)

Par Frédéric André et Didier Moyart de l' ARCSI

Frédéric André est un des fondateurs de Calypt, une société basée sur Lyon et spécialisée dans la sécurité de l'information. Il y réalise notamment des audits de sécurité « offensifs ». Il est passionné par la cryptologie et son histoire.

Didier Moyart vient de l'univers de la carte à puce et travaille pour ALSTOM dans le domaine des télécom.

L'ARCSI, Association des Réservistes du Chiffre et de la Sécurité de l'Information, est une association fondée en 1928 pour accueillir en son sein les réservistes de l'armée de terre. Elle a accueilli successivement tous les réservistes quelque soit le corps d'armée, les employés d'ambassade, puis plus récemment des personnes qui ont travaillé ou sont encore actifs dans le domaine de la sécurité de l'information.

Terminologie

La Cryptologie ou « science des secrets » englobe deux disciplines : la Cryptographie et par extension la Cryptanalyse.

Le mot **Cryptographie** provient étymologiquement de **:** kruptos (κρυπτός) « caché » et graphein (γράφειν) « écrire ». Le but de la cryptographie est de protéger les messages en assurant la confidentialité, l'intégrité et l'authenticité. Pour assurer le secret des messages, ces derniers sont **codés** ou **chiffrés** (ou encore « cryptés » ; terme marketing improbable).

Déchiffrement : opération inverse du chiffrement qui consiste à obtenir la version originale (« clair ») d'un message qui a été précédemment chiffré, en connaissant la méthode de chiffrement et la clef utilisée.

La Cryptanalyse est la science qui consiste à obtenir le clair d'un message qui a été préalablement chiffré, sans connaître la clef de chiffrement ou encore éventuellement la technique employée. On parle alors de décryptement.

Histoire de la cryptologie

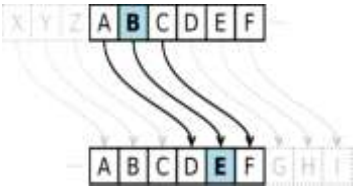
Depuis très longtemps et l'invention de l'écriture les Sumériens 3600 av JC, l'homme s'est ingénié à protéger le secret de ses écrits.

***L'antiquité et les balbutiements de la cryptographie**

- Les égyptiens dès 2300 av JC utilisaient déjà des codes, encore

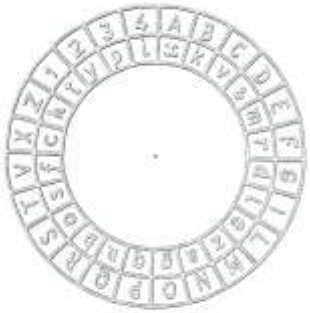
que ce ne fût pas, semble-t-il pour garder un secret mais plutôt pour aiguiser la curiosité.
- Chez les grecs, vers 400 av JC, l'acheminement discret de l'information s'effectuait de la manière suivante: le texte clair était écrit sur une lanière entourée sur un bâton (la **scytale spartiate**) et complété par des caractères aléatoires. Pour lire le message un bâton du même diamètre est nécessaire : la clef du code est le diamètre du bâton

- Chiffre de César. Il consiste en un décalage de 3 lettres de l'alphabet. C'est un système basique simple mais efficace pour l'époque



Le « rot13 » est une variante souvent utilisée aujourd'hui pour les jeux
En généralisant et par substitution d'une lettre par une autre lettre, on obtient 26 ! (factorielle 26) possibilités de substitutions de l'alphabet latin ce qui n'est pas si mal, mais facile à décoder par analyse des fréquences d'apparition des lettres. (en Français : e, s, a, i, t, n, r, ….) Ainsi, si dans un message codé le « t » apparait le plus souvent, c'est qu'il remplace certainement le « e ».

- Au 9ème siècle, percée des arabes qui mettent au point l'analyse des fréquences mais semblent l'avoir utilisé principalement pour des motifs littéraires.
  La solution du chiffrement monoalphabétique mit longtemps avant de se « retrouver » en Europe et entre-temps fut d'ailleurs oubliée en orient.



Cadran d'Alberti

*** Moyen-Age et Renaissance**

- Milieu du XVI^{e} siècle : Alberti propose un cadran de chiffrage sur le principe de substitution
- Et en 1533 Giovani Battista Bellaso
  et Blaise de Vigenère proposent la substitution poly alphabétique : dans le code de Vigenère une lettre est substituée à chaque lettre mais décalée chaque fois dans l'alphabet

  Exemple simple : EFFACEE codé avec la clef "BAC" =>...

  Clair E F F A C E E
  Clef B A C B A C B
  Chiffré F F H B C G F

  On voit que la lettre F n'est pas codée de la même façon.

Vers 1630, Antoine Rossignol invente un système à répertoire code comportant plus de 500 nombres différents. C'est le "Grand Chiffre". Rossignol devient célèbre après avoir fait basculer le siège de la ville Huguenote Réalmont à l'avantage du Prince de Condé, en décryptant un message mettant en évidence les difficultés des assiégés. Il récidivera pendant le siège de La Rochelle.

***A l'aube de la grande Guerre**

Jusqu'au début du XX^{e} siècle, il n'y a pas d'avancée majeure en cryptologie, et un des codes les plus avancés, le Vigenère, a été « cassé » en 1863 par le prussien Kasiski.
On travaillait alors beaucoup avec des répertoires de code dans lesquels les mots étaient remplacés par des groupes de chiffres ce qui permettait le décodage.

***Pendant la Grande Guerre**

- Le fait nouveau est que la radio a été utilisée massivement amenant l'interception des messages cryptés envoyés (le volume rendait un décryptage plus facile). Cela n'a pas été sans conséquences sur la conduite de la guerre

Quelques exemples :

- janvier 1917 : la veille de leur déclaration de guerre, les anglais avaient coupé les câbles transatlantiques issus de l'Allemagne, permettant ainsi l'écoute de leur trafic. Cela permettra l'interception et le décryptement par les anglais du « télégramme Zimmermann » envoyé par les allemands à leur ambassadeur au Mexique, qui annonçait l'extension de la guerre sous-marine et une demande au Mexique d'attaquer les USA. Publié à l'époque de

façon tronquée, et sans indiquer le moyen d'interception et de décodage il a provoqué l'indignation de l'opinion publique aux USA déjà motivée par le torpillage du Lusitania et provoqué l'entrée en guerre des USA.

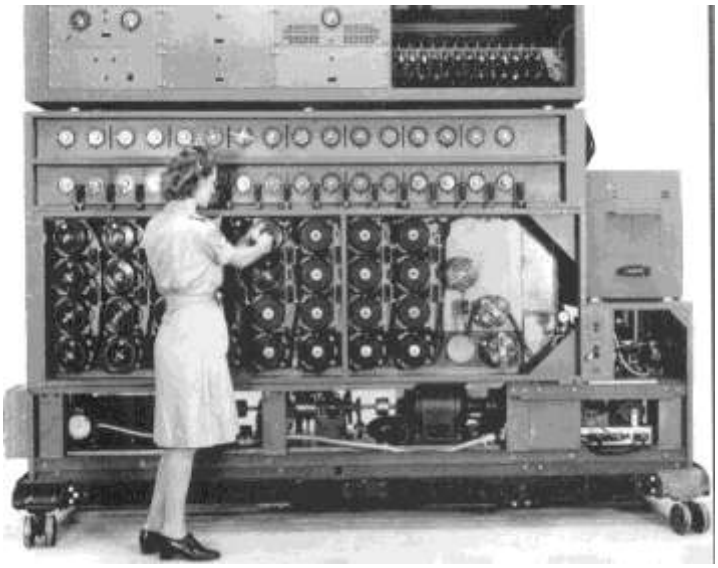
- Le 1er juin 1918 George Painvin décrypte le "Radiogramme de la Victoire". Avant l'arrivée des troupes américaines, la France est en grande difficulté et fait face à une offensive allemande majeure. Painvin parvient à casser le code allemand et traduit alors un message suivant : « Hâtez l'approvisionnement en munitions, le faire même de jour tant qu'on n'est pas vu ». Le message fut transmis au quartier général de Foch, qui fut convaincu de l'imminence de l'attaque sur Compiègne. Les dernières troupes de réserve furent placées autour de la ville et repoussèrent l'attaque. L'Allemagne n'aura alors plus jamais la main dans le conflit.

⁎De l'après guerre de 1914/18 à la guerre de 1939 /1945 :

A la fin de la guerre de 1914, on a appris que les codes « classiques » étaient cassés, et on a cherché à mécaniser le codage pour le rendre plus robuste.
Dès 1917, de nombreuses machines mécaniques et électromécaniques voient le jour (Hebern, Enigma, Hagelin…) et un plus tard, les machines Enigma emblème des problèmes de codage et de décryptage pendant la guerre 40/45. De nombreux textes ou film ont été publiés sur le sujet ; Pour résumer :

La machine Enigma :

Ce type de machine est issu d'un brevet datant de 1926 déposé par Scherbius. Les premières applications militaires datent de 1928 date à laquelle les Polonais interceptent les premiers messages 1928 (c'étaient les plus motivés, ayant le plus à perdre lors d'un conflit).
En 1938, une équipe de mathématiciens polonais parvient à lire 75% du trafic Enigma. La machine est connue, le renseignement français ayant obtenu plans et procédures d'utilisation.
Après les défaites polonaise de 1939 et françaises de 1940, les savoirs des services du chiffre polonais sont transmis à l'Angleterre.
Le GCHQ (littéralement « Quartier général des communications du gouvernement ») officiellement sans existence, a tout mis en œuvre pour briser les codes d'Enigma. Des équipes dont Alan Turing a fait partie ont mis au point des dispositifs électromécaniques appelés « bombes » permettant un décodage journalier des messages. Ces machines travaillaient en parallèle à la recherche de mots courants pouvant être présents dans les messages. Dès qu'un mot en clair était trouvé, les machines s'arrêtaient, la configuration de machine notée et les messages pouvaient être décodés.



En fait des erreurs de protocole et des manies d'opérateurs ont facilité la tache des décodeurs.

Le principe des machines Enigma :

Le principe de base des machines Enigma repose sur l'utilisation de rotors qui transforment l'alphabet clair (noté en minuscules) en alphabet chiffré (en majuscules). Pour mieux l'illustrer, nous nous limiterons à un alphabet de six lettres. Voici la représentation de l'un de ces fameux rotors, ainsi que le schéma équivalent qui permet de mieux suivre l'opération "avec les doigts".

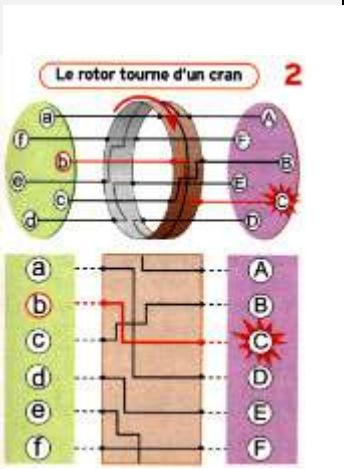
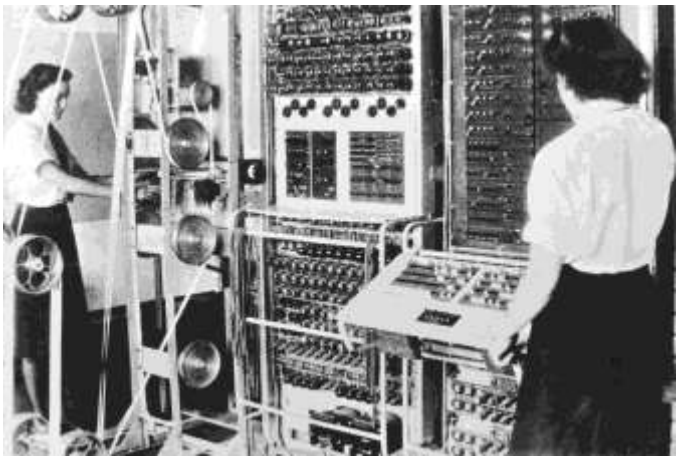


Schéma 1
Si on frappe la lettre **b** sur le clavier, un courant électrique est envoyé dans le rotor, suit la câblage interne, puis ressort à droite pour allumer la lettre **A** sur le tableau lumineux. **b** est donc chiffré en **A**. Idem pour les cinq autres lettres: **a** devient **B**, **b** devient **A**, **c** devient **D**, **d** devient **F**, **e** devient **E** et **f** devient **C**.
Autre principe de base: chaque fois qu'une lettre est tapée au clavier, le rotor tourne d'un cran. Ainsi, **b** devient **A** la première fois, mais **b** devient **C** la deuxième fois (voir schéma 2, à droite), puis **b** devient **E**, etc.
Schéma 2
Dans notre exemple le mot **bac** est chiffré **ADD** (et non **ABD** si le rotor était resté immobile). Pour augmenter le nombre de combinaisons possibles - et déjouer les tentatives des cryptanalystes -, Scherbius a associé plusieurs dispositifs, comme indiqué sur le schéma 3.
Chaque jour la machine avait une position particulière pour les rotors et à chaque message les configurations changent.

- Enigma n'a pas été le seul moyen de télécommunication codé qui a été cassé. C'est le cas des messages assurés par téléimprimeur sécurisé (télex). Les allemands ont utilisé le Siemens T-52 Geheimschreiber , Lorenz SZ-40, et plus tard, le T-43 Siemens.. Le décryptement des messages produits par la Lorenz SZ a nécessité la construction du premier ordinateur ! Le Colossus détruit volontairement après guerre, puis reconstruit (aujourd'hui au musée de l'informatique à Bletchley Park).



L'ensemble des informations recueillies par décryptage étaient appelées ULTRA . Elles étaient sensées provenir de l'espionnage normal car il était impératif que les « écoutés » ne se doutent de rien.
Churchill a déclaré : C'est grâce à ULTRA que nous avons gagné la guerre ! (on pense que ce travail a réduit la durée de la guerre de deux ans)

L'ère électronique

On peut rappeler que le transistor date de 1947 et le circuit intégré de 1958.
Après la guerre le Chiffre français est en déliquescence. Lors du conflit de Suez en 1956, elle s'est aperçue que ses messages n'avaient aucune confidentialité pour ses alliés.
Dans un sursaut d'orgueil, la France se lance dans des études et en 1963 : lors d'un concours de l'OTAN, les français proposent la MYOSOTIS première machine à chiffrer entièrement électronique basée sur des transistors.

Problématiques actuelles

Les principes de la cryptologie :

La crypto est utilisée couramment dans la vie de tous les jours sans qu'on le sache :

communications sur Internet, paiement par CB , monnaie électronique , passeport, chiffrement de fichier, téléphonie mobile, stockage de mots de passe signature électronique...

Conventions de cryptologie : Dans les articles de cryptologie apparaissent des personnages conventionnels

Alice et Bob sont les correspondants , Eve l'écouteuse est la personne qui cherche à intercepter, Mallory est un attaquant actif, Trudy un intrus, Nestor un tiers de confiance…….

Les 3 buts de la cryptologie

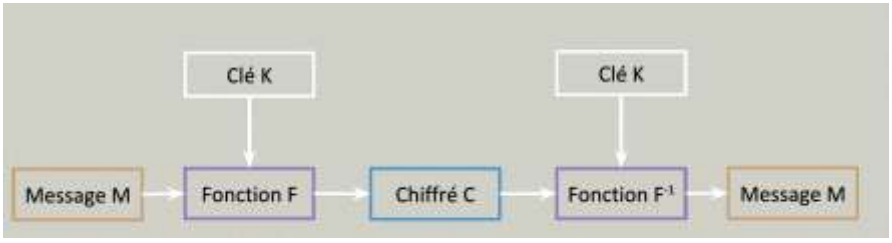
Confidentialité : C'est le fait d'assurer que l'information n'est accessible qu'à ceux à qui elle est destinée.

Authenticité : L'authenticité est l'assurance qu'un message, provient bien de la source / personne dont il prétend venir

Intégrité : L'intégrité signifie qu'une information n'est pas modifiée lorsqu'elle est transmise donc que le message reçu n'est pas altéré.

∗Confidentialité , Les 2 modes de cryptage

Le mode symétrique



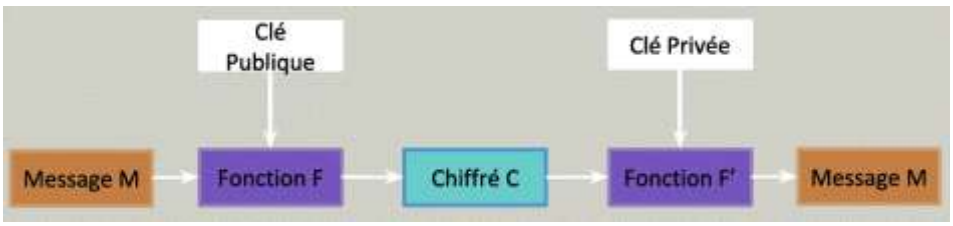
Le message est codé au moyen d'une fonction de mathématique associée à une clef. Le décodage est assuré par la fonction mathématique inverse associée à la même clef. La taille des clefs et les temps de calcul restent raisonnables pour une transmission entre 2 correspondants, mais le nombre de clefs croit exponentiellement avec le nombre de correspondants.

Les algorithmes des systèmes les plus courants comme le DES (Data Encryption Standard) créé en 1977 (clefs de 56 bits) et son remplaçant, l'AES (Advanced Encryption Standard) depuis 2001 utilisent des clefs de 128 à 256 bits. Les algorithmes utilisés sont publics, donc connus de tous mais les clefs doivent être secrètes.

Un peu de statistiques :

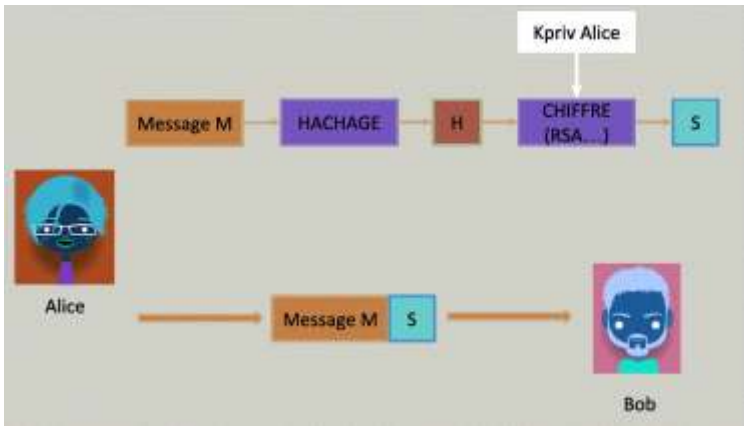
- 1 chance sur 1,4 10^{7} de gagner au loto
- 1 clé sur 7,2 10^{16} pour les combinaisons pour le DES
- et 1clé sur 3,4 10^{38} pour l'AES

Le mode asymétrique



La cryptographie asymétrique, ou cryptographie à clé publique, est une méthode de chiffrement. Elle repose sur l'utilisation d'une clé publique (qui est diffusée) et d'une clé privée (gardée secrète), l'une permettant de chiffrer le message et l'autre de le déchiffrer. Ainsi, l'expéditeur peut utiliser la clé publique du destinataire pour chiffrer un message que seul le destinataire (en possession de la clé privée) peut

décoder. Inversement, l'expéditeur peut utiliser sa propre clé privée pour signer un message que le destinataire peut vérifier avec la clé publique de l'expéditeur (voir ci-dessous);
L'in des systèmes les plus utilisées est le RSA (du nom de ses inventeurs Rivest, Shamir, Aldeman) créé en 1977 Comme tous les algorithmes à clé publique, il repose sur une fonction mathématique difficile à réaliser. Dans ce cas précis, il s'agit de la factorisation des grands nombres. L'avantage de ce type de chiffrement est qu'il ne nécessite qu'un couple de clefs par utilisateur, l'inconvénient est que la taille des clefs peut aller jusqu'à 4096 bits (pour le RSA) et que le décodage nécessite des temps de calculs importants.
En pratique cette méthode permet de transmettre de façon sure une clef symétrique utilisée alors par les 2 correspondants.
On tend actuellement à utiliser des courbes elliptiques ou logarithmiques qui utilisent des clefs plus petites et donc des temps de traitement plus court).

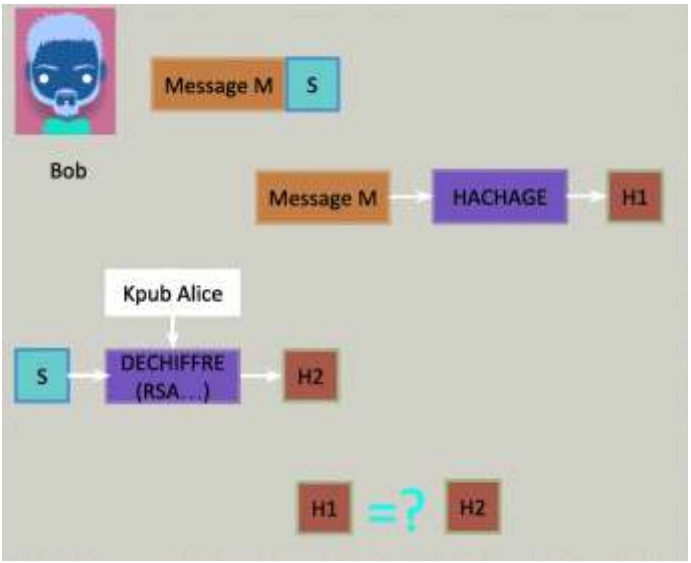


∗L'intégrité des messages :

Elle est réalisée par une fonction de hachage qui, à partir d'une donnée fournie en entrée, calcule une *empreinte* servant à identifier la donnée initiale. Un message traité va fournir une chaine de caractères avec des spécifications bien définies et donnant des informations sur le message sans donner accès au contenu de ce dernier. Le résultat du hachage est appelé « hash » ou « empreinte » C'est une fonction à sens unique qui ne permet pas de remonter au message initial, et qui doit résister aux collisions : la moindre modification du document entraîne la modification de son haché (par exemple l'empreinte d'un chèque de 100 € à Mme Libellule doit être différente de l'empreinte d'un chèque de 1000 € bien que la différence ne soit que d'un caractère) .
Lors de la réception du message, il suffit au destinataire de calculer le hash du message reçu et de le comparer avec le hash accompagnant le document. Si le message ou le hash a été falsifié durant la communication, les deux empreintes ne correspondront pas.

∗L'authenticité d'un message :

Il faut pouvoir prouver que le message a bien été envoyé par Alice et non par un intrus, et que le message ne peut plus être modifié.



Alice hache le message qu'elle doit signer et chiffre le hash résultant avec sa clef privée et obtient un « S » (sceau) qu'elle transmet à Bob avec le message.
A réception du message, Bob le destinataire déchiffre le sceau avec la clé publique de l'expéditeur, puis compare le haché obtenu H2 avec le haché H1 reçu en pièce jointe

<<<<<<< Vérification de signature

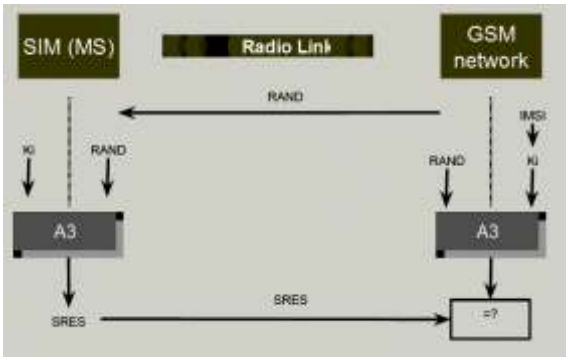
Les applications de tous les jours :

Pour assurer une transaction sécurisée sur Internet, on passe par une notion de certificat délivré par un tiers de confiance. C'est ce qui permet d'avoir des échanges chiffrés entre le navigateur internet et le serveur, comme dans les contacts avec un site de vente.

Le processus nécessite 4 échanges avant de pouvoir établir une liaison codée entre le serveur et le navigateur : 1- Le navigateur demande un certificat au serveur, 2- celui-ci lui transmet ce certificat contenant sa clef publique. 3- Après vérification de la non altération du certificat, le navigateur transmet, chiffré avec la clef publique du serveur, une clef de session. 4- le serveur déchiffre la clé et ouvre la session. Les 2 correspondantont alors chacun la même clef et peuvent communiquer de façon sécurisée. (Le cadenas apparait alors fermé dans le coin gauche de la ligne URL du navigateur) . Ici on retrouve une application du chiffrement assymétrique et symétrique. Le certificat utilise des clés publiques qui sont comme on l'a vu ci-dessus longue à déchiffrer. On négocie ensuite une clé symétrique qui sera utilisée pour le reste de la transaction.

Application GSM

Le mécanisme d'authentification du terminal caractérisé par sa carte SIM est assez analogue au système décrit précédemment. Chaque carte SIM possède une clef (K1) qui lui est propre et qui est connue au centre d'authentification du réseau. L'authentification s'effectue par comparaison entre les résultats d'un calcul sur un nombre aléatoire RAND effectué avec le même algorithme dans la carte SIM et dans le centre d'authentification. Si les résultats concordent, l'utilisateur est reconnu et accepté par le réseau.



Les écueils de la cryptographie

∗Toute la cryptographie moderne symétrique repose sur des algorithmes et des schémas complexes et la sécurité repose sur la résolution/non résolution de problèmes complexes.

Les algorithmes de chiffrement reposent sur quelques familles de fonctions à sens unique. : Faciles de chiffrer, difficiles de déchiffrer sans clef. Ainsi le système RSA repose sur la factorisation de grands nombres : Il est facile de multiplier des nombres premiers , très difficile de factoriser un nombre ayant beaucoup de chiffres.

De même il existe des systèmes utilisant les logarithmes ou la résolution de courbes elliptiques.

∗Mais, si des avancées importantes se faisaient jour dans la résolution de problèmes mathématiques, cela mettrait à mal les systèmes de cryptographie. Comme il se produit des avancées régulières, on peut dire que la sécurité est une notion ponctuelle.

∗L'algorithme de chiffrement offrant une sécurité absolue existe : le chiffrement par la méthode du masque jetable ; (Vernam, 1917). Il consiste à combiner le message en clair avec une clé aléatoire de même longueur à n'utiliser qu'une seule fois.

∗La sécurité des modules cryptographiques repose également sur de nombreux facteurs
Leur implémentation, le câblage, la façon d'utiliser les clefs ! - générer une clé aléatoire n'est pas si facile, (on a encore des stations radio décamétriques qui transmettent des suites de chiffres aléatoires) .Transmettre les clés de façon sécurisée est compliqué. (Dans certains films, on montre une personne qui transporte des codes en ayant une valise menotée à un bras)…

∗La plupart des produits commercialisés contiennent des « backdoor » permettant un décryptage discret. " (La NSA National Security Agency avait pendant longtemps inséré une porte dérobée dans tous les outils logiciels de chiffrement des données). Exemple de Crypto AG société suisse grand fournisseur de système de chiffrement après la Seconde Guerre Mondiale. Les systèmes de cryptage vendus par CryptoAG étaient trafiqué par la NSA de façon à ce qu'elle livre la clef utilisée (choisie par l'utilisateur de la machine et donc inconnue de tous) lors de l'interception du message! De plus, la clef était codée avec un code différent uniquement connu par la NSA pour éviter que cela ne soit trop transparent aux utilisateurs avertis...

∗Aujourd'hui, « Bullrun » est un programme américain secret, utilisé par la NSA, ayant entre autres, pour but de casser des systèmes de chiffrement. L'existence de ce programme a été révélée en septembre 2013 par Edward Snowden.

∗Des produits de sécurité courants contiennent également des bugs plus ou moins volontaires comme la vulnérabilité « Heartbleed » sur OpenSSL qui traite des échanges sécurisés sur le web, ou « goto fail » sur OS X, erreur assez grossière qui a fait que d'aucuns se sont demandé si elle n'était pas volontaire…).

Bruce Schneier, né le 15 janvier 1963 à New York, est un célèbre cryptologue, un spécialiste en sécurité informatique et un écrivain américain. IL dit : Cryptography plays a role in computer security, but buggy computer systems and vulnerable communications are a reality that cryptography has not solved. Dans son livre « Cryptologie appliquée » il insiste sur les risques liés à des protocoles qui ne seraient pas suivis à la lettre (par exemple, clés envoyées en clair dans des courriels), de même que les défauts inhérents aux ordinateurs (plantage, bugs, complexité trop grande, etc.).

Et pour terminer :

un peu d'humour

13 avril 2015 Transcription par HT, validée par les présentateurs.