Xavier
BURTIN
Master 1 Informatique
Option Cryptologie

Techniques analogiques et numériques de sécurisation des transmissions sonores

Sommaire :

Introduction

Depuis longtemps, on cherche à sécuriser la transmission de données afin de pouvoir communiquer uniquement avec les réels destinataires du message, à partir du moment où il y a un risque d'interception par une tierce personne. On a longtemps utilisé différentes méthodes pour « coder » ces messages et les rendre inintelligible pour toute personne ne connaissant pas la méthode de codage, ou la clef de cryptage utilisée. L'apparition de nouvelles technologies a permis de communiquer avec des personnes de plus en plus loin, et de multiples façons. Il a donc fallu adapter ces méthodes afin de respecter les contraintes techniques et physiques de ces nouveaux vecteurs d'informations. La méthode de transmission de données la plus simple et la plus primitive est la parole. Pourtant, il a fallu l'invention d'un appareil permettant de transporter les signaux sonores pour permettre le brouillage de ces ondes analogiques et obtenir ainsi une méthode de cryptage. L'avancée technologique a permis de convertir ces signaux en signaux numériques, ce qui permet encore un autre niveau de cryptage.

Il est question dans ce rapport, de présenter les techniques de base permettant de sécuriser des communications vocales, puis d'exposer les différentes méthodes de conversion de signal analogique en signal numérique, qui est l'étape primordiale précédant le cryptage de ces données.

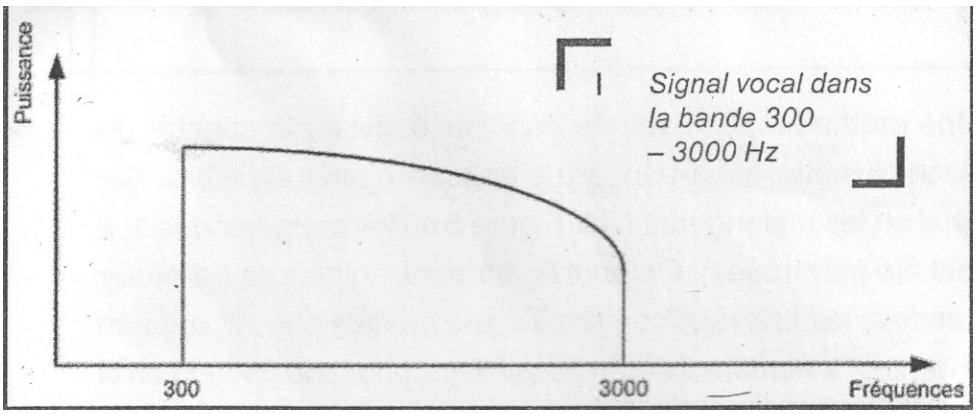
1. Techniques analogiques

Un son est une onde très complexe qui est entendue par l'oreille humaine et interprétée afin d'en restituer un sens. Il est donc tout à fait possible pour une oreille habituée d'entendre et de comprendre la transmission d'un son, et ce même si l'onde est brouillée. De plus, la capacité de l'oreille et du cerveau à s'adapter augmente les chances de saisir des informations sur un canal brouillé au bout de plusieurs écoutes. Un signal analogique brouillé reste un signal continu trop complexe pour pouvoir être formalisé. La voix contient des informations sur le message, mais aussi concernant son auteur (intonations, façon de parler, langage utilisé...). Il est donc difficile de crypter les informations uniquement par des algorithmes. L'intervention d'une ressource humaine paraît alors nécessaire.

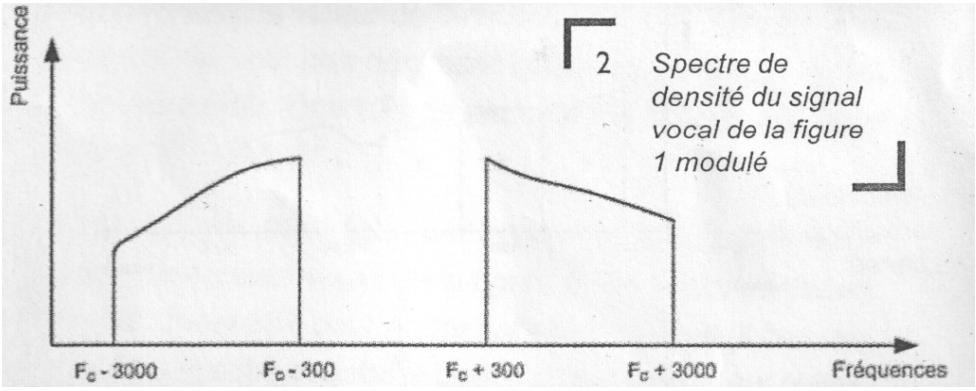
1.1. Techniques d'inversion

Dans un système de communication, le signal sonore est constitué d'une fréquence fondamentale et de fréquences harmoniques. Il est possible de le représenter à l'aide de son spectre de densité, qui montre l'évolution de sa puissance en fonction de sa fréquence. C'est sous cette forme que le signal sera représenté par la suite.

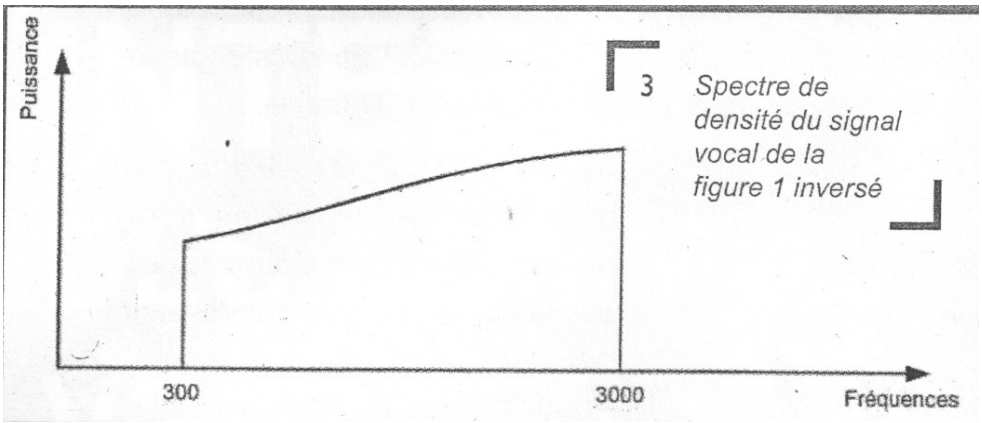
Ces techniques reposent sur le principe d'inversions de fréquences. Le dispositif utilisé est appelé un *inverseur de voix*. Cet inverseur intervertit les hautes et les basses fréquences du signal. On introduit ensuite un second signal appelé modulateur équilibré afin de supprimer la fréquence porteuse du signal. On prend comme exemple le signal suivant :



Soit le signal d'origine d'équation Vm*cos(ωm), et le signal injecté Vc*cos(ωc). V représente l'amplitude du signal, et ω la fréquence. Vc et ωc sont choisis librement (Vc=1, ωc>ωm).L'équation du signal obtenu après injection du signal modulateur est alors : 1/2*(Vc.Vm*cos(ωc+ωm))*t + 1/2*(Vc.Vm*cos(ωc-ωm))*t.On a donc le signal suivant :



On a une fréquence fc = ωc/(2Π). Cette fréquence est appelée la fréquence porteuse. Elle a pour effet de translater le signal dans le domaine fréquentiel. Après filtrage sur la zone de fréquence du signal, on obtient un signal dont les hautes fréquences et les basses fréquences ont étés échangées :



Ce dispositif est purement technique et n'est pas en lui-même un dispositif de cryptage. On

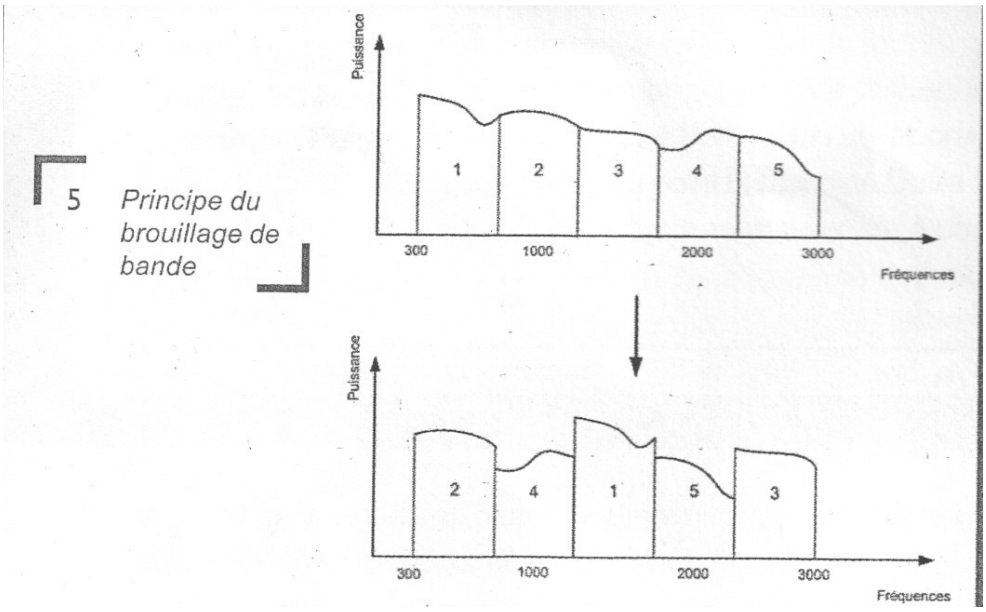
peut facilement retrouver le signal d'origine par des techniques de traitement du signal. Pour insérer un élément de sécurité, il faut utiliser un *inverseur à décalage de bande*, qui va utiliser des fréquences porteuses variables, et donc servir de clef. Le principe est de décaler le signal d'une certaine fréquence, de « couper » la bande de fréquence au-dessus de la fréquence maximum du signal, et de « recoller » cette bande au début du signal. Sur l'exemple, si on a un choix de fréquences porteuses sur 4000Hz, le signal d'origine (compris entre 300 et 3000Hz) risque de se retrouver entre 1000 et 3700Hz. Il suffit de décaler la bande de 700Hz supérieure à 3000Hz vers 300Hz.

On dispose de 4 à 16 fréquences porteuses différentes. Un générateur aléatoire choisit une de ces fréquences à intervalles réguliers.

Ce système a deux faiblesses. La première est que l'on a pas assez de possibilités pour le choix des fréquences. Il est donc possible de trouver la fréquence utilisée pendant l'intervalle de temps, et de reconstituer le signal. La deuxième faiblesse est que le son n'est pas assez brouillé pour empêcher un décodage par une oreille moyennement entrainée.

1.2. Techniques de brouillage de bande

La méthode suivante consiste à diviser le spectre du signal en plusieurs sous-bandes de largeurs égales et de brouiller le signal en les permutants. Il s'agit de changer la position des sous-bandes dans la bande de fréquence du signal. On obtient alors N! Possibilités. Par exemple, pour N=5, on a 3840 combinaisons possibles. En voici une :



5 *Principe du brouillage de bande*

Cette méthode peut également être combinée avec la méthode précédente pour chacune des sous-bandes. On obtient alors (N!)*(2^N) possibilités.

Malheureusement, toutes ces possibilités ne sont pas exploitables. En effet, certaines combinaisons donnent des positions très proches du signal original. Par exemple, le signal initial (1,2,3,4,5) peut être permuté en (5,2,3,4,1). Seulement 2 sous-bandes ont été interverties. Il vaut mieux considérer les dérangements, c'est à dire lorsque des sous-bandes ont réellement changées de place. On obtient le nombre de dérangements par la formule Nd = (N!)/e. On obtient alors 1408 possibilités. De plus, il se trouve que 40% de l'énergie d'un signal est dispersée dans les 2 premières sous-bandes. Il est donc possible de les retrouver et de réordonner les sous-bandes, ou du moins d'en retrouver le fragment du début du message.

Il est donc nécessaire de passer par une validation du brouillage par l'écoute répétée de toutes ces permutations afin d'être sûr que le signal est bien brouillé. Les meilleures solutions sont mises en mémoire. La mémoire utilisée dépend du nombre de combinaisons possibles. Par l'expérience, on sait que pour 5 sous-bandes, il n'y a que 32 permutations. Grâce à l'inversion, on monte à 32*2¨5=1024 combinaisons. Ces dernières sont donc stockées sur 10 bits. Une des permutations mise en mémoire est alors sélectionnée aléatoirement à intervalle régulier grâce à un générateur aléatoire.

La dernière étape fait appel au talent du manipulateur du système qui doit définir le nombre de sous-bandes et la période du générateur aléatoire. Les restrictions sont un nombre de sous-bandes inférieur ou égal à 20 pour éviter de dégrader le signal, et une période assez courte pour s'assurer que personne ne puisse tester toutes les combinaisons.

Cette technique n'assure cependant qu'un niveau de sécurité relativement faible. En effet, la possibilité de décoder le message par l'oreille humaine est non négligeable. Ce dispositif est donc utilisé pour des trafics peu sensibles.

1.3. Techniques de multiplexage temporel

Pour utiliser cette technique, il faut tout d'abord divisé le signal en plusieurs trames temporelles égales, puis diviser ces trames en plusieurs segments de même durées. Les éléments de sécurité sont la durée de chaque segment, et la permutation des segments dans la trame.

Le choix de la durée T dépend de plusieurs facteurs :

- T doit être suffisamment court pour que l'information du segment ne soit pas trop importante (et rendre inutile le brouillage), mais pas trop courte pour ne pas dégrader le signal. On peut aussi utiliser les deux dispositifs précédents pour augmenter la sécurité, mais en faisant attention à ne pas dégrader davantage le signal.
- Le choix de T est imposé par les délais acceptés pour la récupération du message. Si M est le nombre de segments, il faudra M*T secondes pour faire la permutation des segments, et encore M*T secondes pour la permutation inverse. Soit 2*M*T secondes pour la transmission de la trame.
- La durée d'une trame entière (M*T) intervient sur le niveau de sécurité du message. Plus la trame est courte, et moins il y a d'information véhiculée dans chaque trame. La trame est donc plus facile à intercepter.

Le choix de la durée de T est donc laissée au libre arbitre de l'utilisateur du dispositif qui doit trouver les meilleurs paramètres pour la sécurité et l'efficacité de la transmission. Encore une fois, les résultats obtenus doivent être validés par des tests d'écoute. Les valeurs retenues généralement sont de 8 à 24 segments, et une durée T de 20 à 60 ms.

Le deuxième facteur important dans la sécurité d'un système à multiplexage temporel est la permutation des segments dans la trame.

Chaque trame possède sa propre combinaison de permutations, choisie grâce à une clef de cryptage. Le problème qui se pose est exactement le même que dans le dispositif de brouillage de bandes. Le nombre de permutations est N! (N le nombre de segments), mais les permutations les plus proches de la trame initiale sont à exclure. Il faut donc ne considérer que les dérangements et ne stocker en mémoire que les permutations acceptables. De plus, il faut que le signal soit le plus distordu possible. On mesure cette distorsion à l'aide d'un indice appelé le dérangement moyen D. Il se calcule par la formule : $\frac{1}{M} * \sum^{M} |i - \Pi(i)|$.

L'intelligibilité du signal est inversement proportionnelle à la valeur de dérangement moyen. Il est nécessaire d'éviter certaines configurations qui laisseraient deux segments consécutifs trop près de l'autre (deux segments consécutifs ne doivent pas être à moins de trois positions l'un de l'autre). On ignore donc toutes ces permutations.

Enfin, seule la permutation définie par la clef doit être en mesure de restaurer la trame originelle. En effet, aucune autre des permutations en mémoire ne doit être capable de restaurer un un signal qui serait proche du signal d'origine. Cette sécurité est mise en place pour éviter qu'un attaquant connaissant le système ne puisse quand même y accéder.

Au final, quelque soient les paramètres choisis pour ce type de dispositifs, les solutions doivent être testées et validés par des tests d'écoute. Le choix de ces paramètres fournis un tel éventail de possibilités, que le niveau de sécurité dépend énormément de la décision personnelle de l'utilisateur.

Jusqu'à présent, les techniques présentées consistaient uniquement à effectuer une déformation du signal analogique afin de brouiller ce dernier. Avec l'apparition du de l'ère numérique, on a tenté de coder ce signal analogique afin de le traiter numériquement. Il a été nécessaire de convertir ce signal en un signal numérique le plus fidèle possible. Le cryptage des données numériques devient alors évident puisqu'il s'agit de crypter des données informatiques. Ce cryptage permet donc un niveau de sécurité plus important que les techniques analogiques.

Les techniques qui suivent décrivent donc la partie critique du processus, à savoir la conversion du signal analogique en signal numérique.

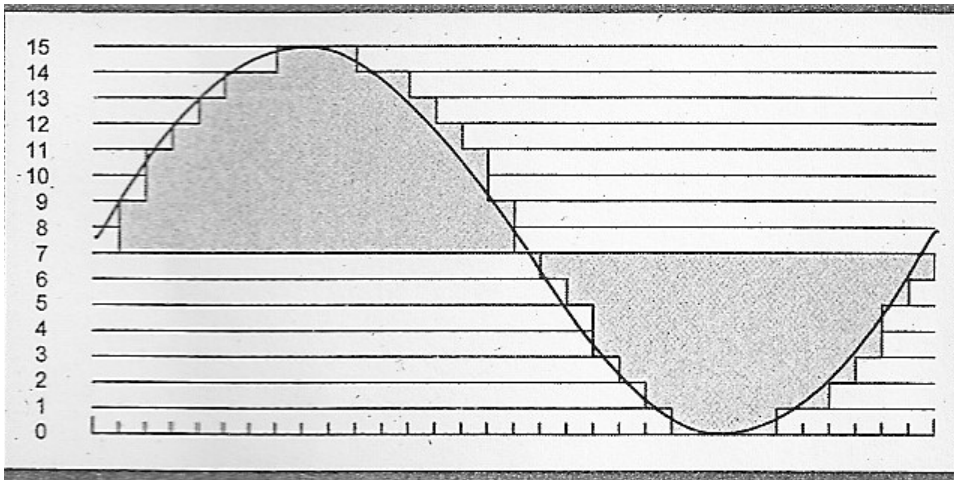
2. Techniques numériques.

Un signal numérique est un signal discrétisé contrairement au signal analogique continu. Il est représenté par une série de valeurs numériques cryptables.

2.1 Modulation par impulsion codée (Pulse Code Modulation : PCM)

Cette technique a été mise au point par Alec Reeve en 1937, et utilisée dans le système SIGSALY, un système de cryptophonie de haut niveau de sécurité des forces alliées. Elle est utilisée de nos jours par certains systèmes téléphoniques, dans le traitement de la voix, par les CDs, et dans le codage vidéo.

Le principe est de considérer le signal comme une série de pulsations qui vont être codées en valeurs numériques. Le signal est échantillonné en intervalles de temps réguliers. La fonction du temps F(t) peut être évaluée à partir de différents paramètres du signal analogique (amplitude, valeur du spectre de densité,...). On transmet la valeur du paramètre choisi au temps t afin d'obtenir des valeurs discrètes. On obtient un signal échantillonné :



Le temps t choisi doit être suffisamment court pour permettre une approche suffisamment précise du signal analogique, et permettre une reconversion du signal numérique en signal analogique efficace. Il faut cependant que ce temps n'engendre pas un nombre trop importants de valeurs à transmettre afin de ne pas surcharger le débit de transmission. Il s'agit de déterminer le meilleur compromis qualité/taux. Le nombre de valeurs par unité de temps est appelée « taux d'échantillonnage ».

Le théorème d'échantillonnage précise que pour une fonction F(t) limitée en fréquences entre les fréquences f1 et f2 Hz (f2-f1=B), le taux d'échantillonnage minimum pour caractériser la fonction F(t) est de 2*f2/M échantillons par secondes, M étant le plus grand entier inférieur ou égal à f2/B.

A partir de ce théorème, on en déduit qu'un signal peut être décrit par 2*B échantillons par seconde (2*f2/M <= 2*B). Les normes considèrent la valeur B à 4000Hz, soit un taux d'échantillonnage de 8000 échantillons par secondes.

On effectue cette conversion en quatre étapes :

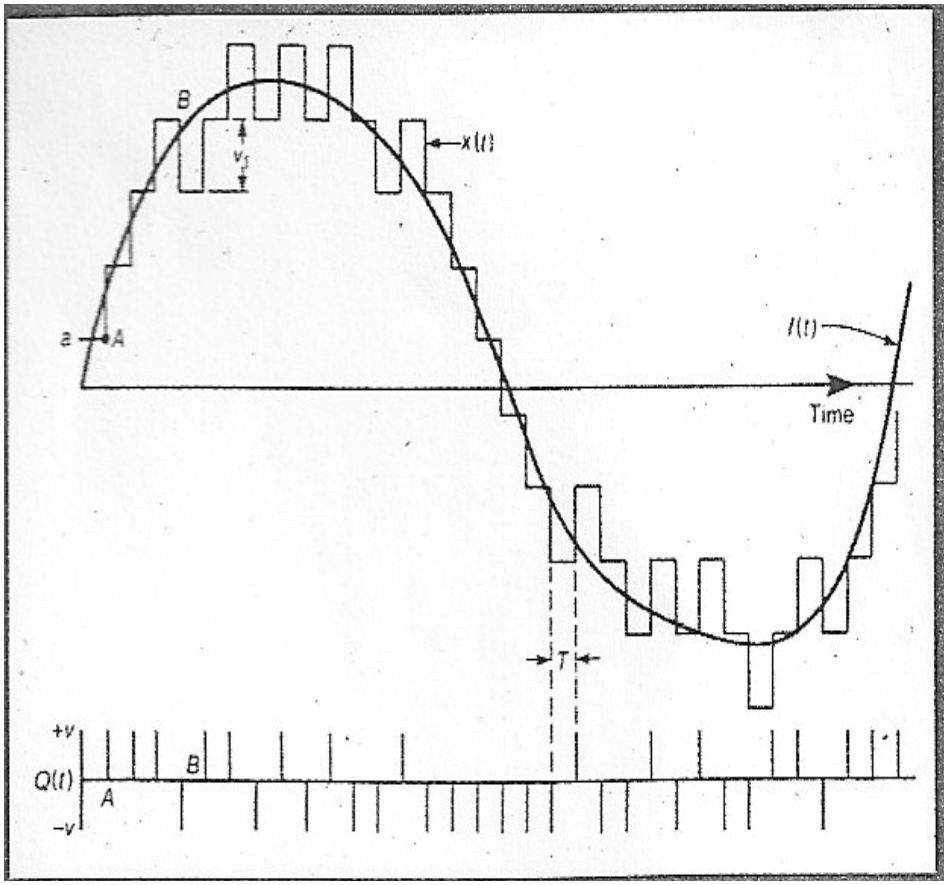
- Etape de filtrage : On limite le signal à la bande de fréquence B.
- Etape d'échantillonnage : On récupère toutes les échantillons de F(t).
- Etape de quantification : Chaque échantillon a une valeur comprise entre -127 et 127, soit 256 valeurs possibles. Coder ces 256 valeurs nécessite 8 bits, ce qui représente une quantité d'information assez importante. On choisit donc une valeur arbitraire la plus proche possible de la valeur réelle de l'échantillon parmi une série de valeurs prédéfinies. Il existe deux lois pour ces valeurs, la loi μ valide aux Etats-Unis et au Japon, et la loi A valide en Europe.
- Etape de codage : La valeur du signal est ensuite traduite en une séquence de 8 bits. Il est possible d'appliquer des corrections d'erreurs sur ces séquences pour limiter le bruit et l'approximation faite.

Il existe trois faiblesses à cette méthode. elle est très gourmande en terme de débit (8000 échantillons par secondes et 8 bits par échantillon nécessitent un débit minimum de 64 Kbits/sec), surtout que certaines variantes nécessitent un codage sur 24 bits. De plus, l'erreur de quantification peut apporter des difficultés de restitution du message. Enfin, il est nécessaire d'avoir une très bonne synchronisation pour restituer le signal analogique.

2.2. Modulation linéaire delta (MLD)

Le principe de cette méthode est de construire un signal en escalier à partir de la valeur du signal échantillonné et la valeur réelle du signal. Le but est ainsi de réduire la quantité de données

transmises. On décrit une fonction Q(t) à partir d'impulsions constantes de valeur v (en volts) à intervalles réguliers T. La fréquence Fp doit être supérieure au taux d'échantillonnage donné par le théorème. On choisit un voltage initial a comme point de départ pour une fonction x(t) pendant la première période T. On compare cette valeur avec la valeur de f(t). Si x(t)<f(t), on augmente x(t) de la valeur v, sinon on la diminue de v. On obtient une fonction en escalier qui approxime la valeur de f(t) :



On obtient alors une série d'impulsions codables numériquement. On ne transmet plus la valeur, mais la variation de Q(t) en fonction de la valeur précédente. On obtient donc un débit de transmission des données bien moindre que dans la technique précédente (environ 9 kbits/sec).

2.3. Vocoders

Contrairement aux deux dernières méthodes qui consistaient à approximer le signal analogique par échantillonnage, cette méthode transforme le signal à l'aide d'une description structurelle et des paramètres de cette description. Par exemple, un signal de forme sinusoïdal est convertit grâce à l'envoi de l'information « signal de forme sinusoïdale », et par les paramètres de cette sinusoïde (amplitude, phase, fréquence...). Le signal est alors reconstruit à partir de ces informations. Un vocoder (voice coder) est donc composé d'un analyseur de signal audio pour déterminer la « forme » du signal, et d'un synthétiseur pour reconstruire le signal à partir des informations transmises. L'économie de bande passante est alors de l'ordre de 90% (2400 bits/sec suffisent).

Il y a cependant quelques défauts :

- Le signal synthétisé n'est pas parfait. Il s'agit d'une reconstruction par ordinateur qui aboutit à une déformation du signal d'origine.

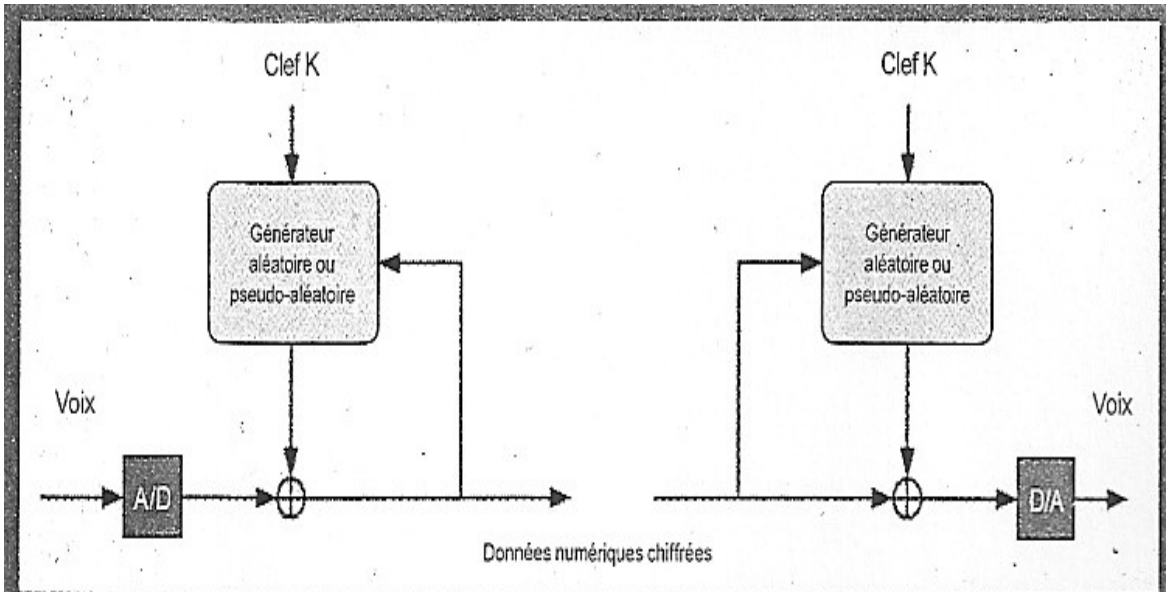
- Le bruit du canal peut engendrer des erreurs importantes sur la transmission du signal. Il faut alors avoir recours à des techniques de réparation d'erreurs. La répétition du signal peut être importante, ce qui augmente au final le débit minimum pour la transmission des données.
- La technique est optimisée pour un groupe linguistique. La façon de parler des européens est différente de celle des asiatiques en termes de sonorités. Un même vocoder n'aura pas la même qualité selon le groupe qui l'utilise.

2.4. Application à la cryptophonie

La technique la plus utilisée est la technique MLD dans les matériels cryptophoniques à cause de son efficacité face aux bruit. Un canal avec un taux d'erreur de 10% peut délivrer un message à un débit de 16 kbits/s. Une fois le signal numérisé, on peut appliquer un système de cryptologie numérique. Il s'agit tout simplement de chiffrer la séquence de la même façon que pour n'importe quel flux de données. En règle générale, on utilise un système par flot de type pseudo-aléatoire. On utilise une clef de 128 à 256 bits pour initialiser le système qui produira la séquence aléatoire combinée à la séquence numérisée.

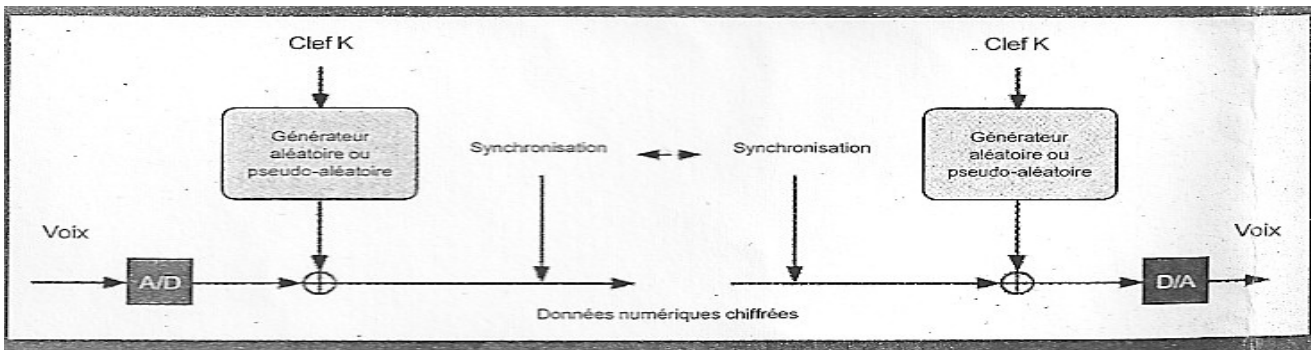
Il existe deux méthodes qui dépendent du type de synchronisation souhaitée :

La synchronisation est nécessaire pour garantir un déchiffrement sans erreur, et pour obtenir une bonne qualité de voix à la restitution.



Le système est appelé « autoclave » sur le chiffré. Le message sert à son propre chiffrement en plus de la clé. Ainsi, il est possible de récupérer toute perte de synchronisation en fonction de l'avancement de la transmission du message. Cela garantit un système plus simple et moins coûteux, en regard d'une sensibilité importante aux erreurs de transmission, qui risquent d'être propagées. Ces systèmes sont donc plutôt réservés à des usages sur de courtes distances.

Le second système supprime cette action du message sur les états internes du systèmes. La synchronisation se fait alors en continue. Elle est aléatoire et transmise à des intervalles de temps aléatoires. Il suffit de mettre en place une clef pour la synchronisation.



Conclusion

On a vu que la sécurisation de la transmission d'un signal sonore passe par une étape importante de traitement du signal. En effet, le signal analogique possède des propriétés physiques extrêmement complexes qui lui son propres. Il est donc impossible de faire une formalisation et une généralisation de la sécurisation des transmissions. La première idée est de brouiller ce signal en agissant sur ces contraintes physiques, mais le résultat n'est efficace que dans certains cas, et il est alors nécessaire de faire intervenir un élément humain dans l'équation. La sécurité résultante n'est pas de très au niveau, mais peut être efficace pour certaines utilisations tactiques à courte durée, car décrypter ses signaux nécessitent aussi une capacité humaine spécifique. Le passage à la numérisation apporte un élément de sécurité beaucoup plus important grâce au cryptage des données numériques. Cependant, elle est très dépendante de la conversion du signal analogique en signal numérique.

Bibliographie

- revue Misc, Num 38, Mai/Juin 2008, Eric Filiol : « Interceptions des communications vocales : Techniques analogiques ».
- revue Misc, Num 38, Juillet/Août 2008, Eric Filiol : « La sécurité des communications vocales : Techniques numériques ».

Sites internet :

- http://fr.wikipedia.org/wiki/Modulation_du_signal
- http://menut.patrick.free.fr/radioamateurisme/techniques/emissionreception.htm
- http://yusynth.net/archives/ETSF/ETSF-CH9.pdf
- http://fr.wikipedia.org/wiki/Cryptographie_sym%C3%A9trique
- http://fr.wikipedia.org/wiki/Chiffrement_par_bloc
- http://fr.wikipedia.org/wiki/Chiffrement_de_flux
- http://www.miscmag.com/index.php/2008/07/01/references-de-l-article-la-securite-des-communications-vocales-3-techniques-numeriques-d-eric-filiol-parut-dans-misc-38
- http://perso.telecom-paristech.fr/~rioul/documents/200604dsp.pdf