STÉGANOGRAPHIE

Stéganographie vient du Grec "steganos" (dissimulé), et "graphy" (écriture) et signifie donc "écriture dissimulée".

En effet c'est l'art de dissimuler des données. Bien que ce qui nous intéresse ici est en rapport avec l'informatique il peut être intéressant de revenir un peu en arrière. Ainsi on se rend compte que la première forme de stéganographie répertoriée nous vient d'une histoire grecque signé Herodotus et datant du 5eme siècle avant Jésus-Christ. L'auteur nous relate la révolte contre les lois Perses. Afin de communiquer secrètement deux chefs de guerre utilisèrent des esclaves. Ils leurs tatouaient sur le crâne le message et ensuite laissaient les cheveux repousser. L'esclave était ensuite envoyé chez le correspondant trompant ainsi l'ennemi. Une fois rasé le message était parfaitement lisible. Bien que très rudimentaire, cette méthode est un assez bon symbole de ce qu'est la stéganographie.

Texte : Le placement des ponctuations, l'introduction de variations orthographiques ou typographiques, le choix entre des synonymes ou des formes grammaticales, l'espacement entre les mots est des façons simples d'ajouter de l'information sans perturber l'information originale. D'autres techniques plus subtiles mais aussi plus délicates à mettre en œuvre, consistent à offrir d'autres clefs de lecture en utilisant que certaines lettres ou les lettres dans un certain ordre. On peut aussi utiliser un générateur de texte 'aléatoire' basé sur une sorte de 'grammaire'. Les choix faits pour la génération du texte correspondent au message secret.

Son : De faibles variations, imperceptible pour l'oreille, dans les basses fréquences ou ce que l'on appelle le bruit de fond peuvent contenir une grande quantité d'information. Un grésillement infime peut cacher des secrets. Evidemment, ce bruit doit de préférence être transmis de façon numérique sans quoi les vraies pertes de transmission pourraient effacer entièrement le message caché. Afin de rester indécelable, le bruit artificiel doit posséder les propriétés statistiques d'un vrai bruit de fond.

Image : De la même manière que pour le son, des variations dans les basses fréquences de l'image peuvent contenir de l'information. Lorsque le transfert d'information est numérique, d'autres détails du codage de l'image, telle la palette de couleur, peuvent également contenir des informations. Pour le 'watermarking' on préférera reproduire plusieurs fois la signature dans les zones les plus contrastées pour que celle-ci résiste le plus possible à des modifications de l'image.

Autres : Naturellement, il existe autant d'endroits où cacher de l'information qu'il existe de formats et de types de données, les plus fréquemment utilisés étant les plus anodins. On peut penser facilement aux formats multimédias, mais aussi aux formats de compression, d'archive et d'encodage. Même une simple page HTML peut contenir plus que ce vous ne pouvez y voir à travers votre outil de navigation.

Exemple sur une image couleur :

Image porteuse

Image à cachée

Image codée (- 5 LSB)

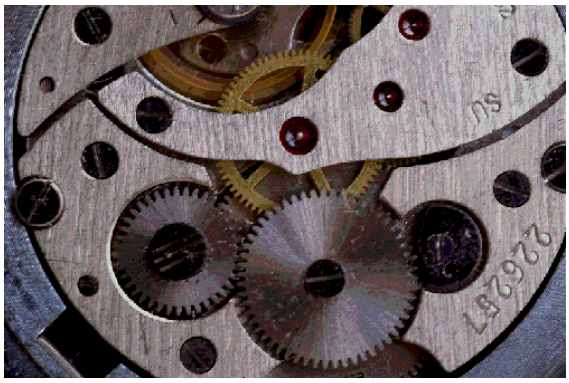

Image récupérée