Stéganographie

Cet article concerne les techniques de dissimulation d'information. Pour l'écriture abrégée, voir Sténographie.

La **stéganographie** est l'art de la dissimulation : son objet est de faire passer inaperçu un message dans un autre message. Elle se distingue de la cryptographie, « art du secret », qui cherche à rendre un message inintelligible à autre que qui-de-droit. Pour prendre une métaphore, la stéganographie consisterait à enterrer son argent dans son jardin là où la cryptographie consisterait à l'enfermer dans un coffre-fort — cela dit, rien n'empêche de combiner les deux techniques, de même que l'on peut enterrer un coffre dans son jardin.

C'est un mot issu du grec ancien στεγανός / steganós (« étanche ») et γραφή / graphế (« écriture »).

1 Histoire^[1]

Dans son *Enquête*, l'historien grec Hérodote (484-445 av. J.-C.) rapporte ainsi une anecdote qui eut lieu au moment de la seconde guerre médique. En 484 avant l'ère chrétienne, Xerxès I^{er}, fils de Darius, roi des Perses, décide de préparer une armée gigantesque pour envahir la Grèce (Livre VII, 5-19). Quatre ans plus tard, lorsqu'il lance l'offensive, les Grecs sont depuis longtemps au courant de ses intentions. C'est que Démarate, ancien roi de Sparte réfugié auprès de Xerxès, a appris l'existence de ce projet et décide de transmettre l'information à Sparte (Livre VII, 239) :

« il prit une tablette double, en gratta la cire, puis écrivit sur le bois même les projets de Xerxès; ensuite il recouvrit de cire son message: ainsi le porteur d'une tablette vierge ne risquait pas d'ennuis. »

Un autre passage de la même œuvre fait également référence à la stéganographie : au paragraphe 35 du livre V, Histiée incite son gendre Aristagoras, gouverneur de Milet, à se révolter contre son roi, Darius, et pour ce faire,

« il fit raser la tête de son esclave le plus fidèle, lui tatoua son message sur le crâne et attendit que les cheveux eussent repoussé; quand la chevelure fut redevenue normale, il fit partir l'esclave pour Milet. » En Chine, on écrivait le message sur de la soie, qui ensuite était placée dans une petite boule recouverte de cire. Le messager avalait ensuite cette boule.

Dès le I^{er} siècle av. J.-C., Pline l'Ancien décrit comment réaliser de l'encre invisible (ou « encre sympathique »). Les enfants de tous les pays s'amusent à le faire en écrivant avec du lait ou du jus de citron : le passage de la feuille écrite sous une source chaude (fer à repasser chaud, flamme de bougie...) révèle le message.

Durant la Seconde Guerre mondiale, les agents allemands utilisaient la technique du micropoint de Zapp, qui consiste à réduire la photo d'une page en un point d'un millimètre ou même moins. Ce point est ensuite placé dans un texte normal. Le procédé est évoqué dans une aventure de Blake et Mortimer, S.O.S. Météores. Il reçoit aussi une belle illustration dans le film de Claude Pinoteau, Le Silencieux.

Un couple célèbre d'artistes de music-hall des années 1960, Myr et Myroska, communiquait les yeux bandés, en apparence par « transmission de pensée » et, en réalité, par un astucieux procédé stéganographique à base de phrases codées (dont, en particulier, des variantes de la phrase : « Myroska, êtes-vous avec moi ? »).

Le principe alors utilisé est toujours largement repris aujourd'hui réf. nécessaire.

2 Méthodes

Supposons, pour notre exemple, que, durant la Seconde Guerre mondiale, une résistante, Alice, doive envoyer tous les jours le nombre de bateaux en rade de Marseille à son correspondant à Paris, Bob. Ils conviennent qu'Alice enverra tous les jours à Bob les prix moyens de divers fruits observés sur le marché de Marseille. Il faut, bien sûr, qu'un agent ennemi, Oscar,

- ne puisse découvrir le contenu caché;
- ne puisse même savoir qu'un contenu est caché;
- ne puisse empêcher la transmission d'un contenu caché éventuel;
- ne puisse envoyer une fausse information en se faisant passer pour Alice.

2 3 CONTRE-MESURES

2.1 Création d'un contenu ad hoc

Alice peut envoyer un message contenant :

Poires: 0 Cerises: 0 Pommes: 1 Tomates: 3 Courgettes: 2

Bob découvrira qu'il y a, ce jour-là, 132 bateaux.

La technique informatique citée ci-dessous comme *Codage sous forme d'une apparence de spam* s'apparente à cette méthode.

L'avantage de la méthode est qu'Alice pourra envoyer à Bob une information très longue. Toutefois, la méthode ne peut être utilisée qu'une seule fois car Oscar pourra rapidement se rendre compte du procédé.

2.2 Modifications mineures d'un contenu existant

Alice peut envoyer un message contenant :

Poires : 4.00 Cerises : 12.00 Pommes : 5.01 Tomates : 3.23 Courgettes : 10.02

Les techniques informatiques décrites ci-dessous dans les rubriques *Usage des bits de poids faible d'une image (LSB)* et *Modulation fine d'un texte écrit* correspondent à cette technique.

L'avantage de la méthode est qu'Alice pourra envoyer à Bob une information relativement longue. Toutefois, Oscar pourrait comparer les prix transmis avec les prix réels (dans le cas du procédé LSB, faire une comparaison bit à bit), pourrait s'étonner d'une précision superflue, pourrait interdire une trop grande précision. (cf. plus bas : stérilisation)

2.3 Dissimulation dans un élément annexe au contenu

Alice peut, le lundi, envoyer un message contenant :

*Poires: 4 *Cerises: 12 *Pommes: 5 *Tomates: 3 *Courgettes: 10

et, le mardi, dans un ordre différent (Alice étant fantasque), mais avec des prix parfaitement exacts :

*Cerises: 12 *Poires: 3 *Tomates: 3 *Pommes: 6 *Courgettes: 10

Le contenu réel du message est dissimulé dans la variation de l'ordre des fruits par rapport à l'ordre de la veille.

L'inconvénient de la méthode est que le message est relativement limité en taille. Si Alice se limite à 5 fruits, elle peut transmettre chaque jour à Bob une valeur comprise entre 1 et 120 (factorielle de 5). L'avantage réside dans la difficulté pour Oscar de repérer l'existence du procédé stéganographique.

Une technique informatique correspondante consiste à

maintenir une image intacte mais à y incorporer une table des couleurs ou *palette* construite dans un ordre qui paraît arbitraire. Le contenu caché peut être une clef donnant accès à un message plus long. En outre, le contenu doit normalement inclure un procédé (généralement un checksum) permettant de vérifier sa validité. L'image qui sert de vecteur à un contenu caché peut être un extrait d'une image connue mais ne peut jamais être sa reproduction exacte, au risque de permettre par comparaison de révéler l'utilisation d'une technique stéganographique.

3 Contre-mesures

Cet art est à risque mesuré dans la mesure où il s'applique à l'information. Son point faible réside donc dans la transmission et la diffusion de cette information.

Une société qui désire contrer l'usage de la stéganographie essayera d'empêcher, de modifier ou de détruire la transmission, la diffusion ou le message lui-même. Par exemple, en interdisant tous contenus arbitraires, abstraits, interprétables, nuancés, fantaisistes, fantasques, poétiques, etc. Elle imposera le respect de critères formels stricts. Ou, au contraire, s'efforcera, dans le secret, de stériliser toutes les informations (cf. paragraphe sur l'imagerie) à des points clés de la transmission des informations (offices postaux, ...). Ce risque de manipulations est encore plus grand avec l'informatique dans la mesure où les interventions humaines sont moins nombreuses assurant ainsi la discrétion des mesures de coercition et les possibilités d'intervention plus grandes (piratage, cheval de Troie...). La destruction systématique de toute information ou des diffuseurs ou récepteurs est sans doute le plus vieux procédé dont la fiabilité n'est pas assurée (de par sa non-exhaustivité dans la pratique; l'information étant humainement vitale).

Dans l'exemple ci-dessus, elle supprimera l'usage de décimales, imposera un ordre alphabétique, interdira les messages dont le contenu ou la langue ne sont pas compris par un préposé, etc.

De nos jours, la stéganographie peut être utilisée à deux fins distinctes : les communications humaines (humains à humains) et machines (machines à machines). Dans les deux cas de figure, il faut au strict minimum qu'il y ait deux parties : un émetteur et un receveur. Cependant, les deux peuvent ne pas se trouver dans le même « espacetemps ». Autrement dit, rien n'empêche de communiquer une information à un tiers n'existant pas encore. Il n'est donc pas improbable de trouver des messages dissimulés jadis. La problématique des contre-mesures à adopter prend alors une toute autre dimension.

La transmission de l'information étant naturelle et vitale à toute société humaine, il n'est pas envisageable de la détruire intégralement (d'où son efficacité limitée intrinsèquement). En revanche, dans le cas des communications machines, des moyens efficaces et terriblement dangereux existent (tels que le nucléaire via la destruction de tout dispositif électronique par ondes électromagnétiques...). Cette contre-mesure extrémiste mettrait à mal toute la société visée. La stéganographie peut être utilisée comme moyen de coercition dans le cadre de communications machines. Par exemple, les virus informatiques et certaines techniques de piratage peuvent en revêtir une forme. La technique du trojan en est également une.

Si l'on occulte les possibilités extrémistes, le meilleur moyen coercitif reste la modification de toute information transmise entre humains ou machines par interventions discrètes ou radicales et non leur destruction.

4 Contre « contre-mesures »

La coercition a l'inconvénient d'engendrer systématiquement des moyens de la contourner, et ce, sans fin envisageable (de par la nécessité de l'information).

La redondance de l'information, des transmissions ou de la diffusion reste le moyen de lutte le plus simple.

L'autre est de ne pas cacher l'information ou de la noyer. Par exemple, cacher de l'information inutile dans un message ouvert utile : le réflexe étant alors de se focaliser sur l'information cachée plutôt que d'admettre l'évidence du message en clair.

Actuellement, à la vue de la quantité du flot continu d'informations qui inonde nos sociétés modernes, il est mathématiquement impossible d'empêcher l'utilisation de la stéganographie qui a l'avantage de pouvoir revêtir d'innombrables formes cumulatives.

5 Techniques rendues possibles par l'ordinateur

5.1 Message transporté dans une image

5.1.1 Usage des bits de poids faible d'une image

L'idée est de prendre un message et de le modifier de manière aussi discrète que possible afin d'y dissimuler l'information à transmettre. Le message original est le plus souvent une image. La technique de base --- dite LSB pour Least Significant Bit --- consiste à modifier le bit de poids faible des pixels codant l'image : une image numérique est une suite de points, que l'on appelle pixels, et dont on code la couleur à l'aide d'un triplet d'octets, par exemple pour une couleur RGB sur 24 bits. Chaque octet indique l'intensité de la couleur correspondante --- rouge, vert ou bleu (Red Green Blue) --- par un niveau parmi 256. Passer d'un niveau n au niveau immédiatement supérieur (n+1) ou inférieur (n-1) ne modifie que peu la teinte du pixel, or c'est ce que l'on fait en modifiant le bit de poids faible de l'octet.



Image d'un arbre. En ne gardant que les 2 bits les moins significatifs de chaque composante de couleur, on obtient l'image suivante (après éclaircissement)



Image d'un chat extraite de l'image précédente.

Exemple Donnons un exemple, considérons l'image

Chaque entrée de ce tableau représente un pixel couleur, nous avons donc une toute petite image 2×2. Chaque triplet de bits (0 ou 1) code la quantité de l'une des trois couleurs primaires du pixel (une image couleur aura dans presque tous les cas des groupes de 8 bits, appelés octets, mais on n'utilise que 3 bits pour clarifier l'exemple). Le bit le plus à droite de chaque triplet est le fameux bit de poids faible --- LSB. Si on souhaite cacher le message 111 111 101 111, l'image est modifiée de la façon suivante : le bit de poids faible du i^e octet est mis à la valeur du i^e bit du message ; ici on obtient :

D'autres techniques similaires sont possibles. Par exemple, l'encodage du message peut être basé sur le

mode de colorisation TSL (Teinte Saturation Luminance) plutôt que RGB (Red Green Blue / Rouge Vert Bleu). Mais toutes ces techniques ont l'inconvénient d'entrainer une déformation - voire une perte - des informations de l'image et sont facilement détectables soit par comparaison avec l'image originelle, soit par analyse linéaire simple (de la parité par exemple!).

Ces techniques de stéganographie très basiques s'appliquent tout particulièrement au format d'image BMP, format sans compression destructive, avec codage des pixels entrelacé sur 3 octets comme énoncé cidessus 2. Réciproquement, tout procédé de compression-décompression d'images avec pertes ou de redimensionnement de l'image est susceptible de détruire un message stéganographique codé de ces façons. On parle alors de stérilisation. Un pays totalitaire pourrait stériliser à tout hasard toute image BMP entrant ou sortant de son territoire, moyennant les ressources techniques nécessaires.

5.1.2 Manipulation de la palette de couleurs d'une image

Certains formats graphiques tels que GIF ou PNG permettent le stockage des couleurs de l'image par référence à une palette de couleurs insérée dans le même fichier.

Ainsi, au lieu de stocker bleu, blanc, rouge dans une image du drapeau français, on trouve dans un format de fichier la description de l'objet la suite couleur1, couleur2, couleur3 ainsi qu'une palette qui définit que couleur1 est le bleu, couleur2 le blanc et couleur3 le rouge.

La même image peut-être stockée de la façon suivante : couleur2, couleur3, couleur1 avec une palette qui définit que couleur2 est le bleu, couleur3 est le blanc et couleur1 est le rouge.

Ces deux images sont visuellement identiques, mais le stockage de celles-ci est différent. Pour une image contenant 256 couleurs uniques dans sa palette, on a factorielle 256 (256!) façons de stocker cette image. En utilisant un code connu entre l'émetteur et le récepteur de l'image, on peut donc communiquer un message de petite taille (log2(256!)), un peu moins de 1 684 bits) caché dans la permutation des couleurs de la palette de l'image.

5.1.3 Message caché dans les choix de compression d'une image

Cette section doit être **recyclée**. Une réorganisation et une clarification du contenu sont nécessaires. Discutez des points à améliorer en page de discussion.

Tout semble indiquer que l'on ne peut cacher un message dans un format d'image utilisant une compression avec perte. En réalité la plupart des programmes de stéganographie sérieux s'attaquent justement au format JPEG qui utilise ce type de compression.

L'idée n'est pas de cacher une information dans les couleurs ou dans la palette (puisqu'il n'y en a pas) mais dans les choix de compression. En effet, tout algorithme de compression nécessite une succession de choix.

Avec des algorithmes de compression tels que Zip ou Gzip, on peut choisir la puissance de compression. En consommant plus de temps calcul et/ou plus de mémoire pour les opérations intermédiaires, on peut obtenir de meilleurs résultats de compression. Ainsi deux fichiers compressés de tailles différentes peuvent être décompressés en deux fichiers identiques.

La compression dans le format JPEG est double. La première compression consiste à découper l'image en blocs de 8 fois 8 pixels et de transformer ces carrés sous une forme mathématique simplifiée. Cette compression introduit des pertes et la version mathématique peut être légèrement différente du carré original tout en étant visuellement très semblable. Une fois tous les blocs compressés, il faut coder les formes mathématiques en consommant le moins possible d'espace. Cette deuxième compression n'introduit pas de perte et elle est similaire dans les principes à ce que l'on peut retrouver dans Zip ou Gzip. C'est en introduisant dans cette phase des bits d'informations que l'on arrive à transporter un message caché.

Voir l'article détaillé : Tatouage numérique.

5.2 Message transporté dans un texte

5.2.1 Modulation fine d'un texte écrit

Décaler une lettre de quelques pixels ne pose aucun problème sur une imprimante à laser et c'est pratiquement invisible à l'œil nu. En jouant sur les interlettrages d'un texte très long et à raison de deux valeurs d'espacement correspondant à 1 et 0, il est possible de transmettre un message sous forme papier, qui ne révèlera son vrai sens qu'une fois analysé par un scanner ayant une bonne précision.

Historiquement, le procédé fut utilisé dès les années 1970 en utilisant non pas des imprimantes laser, mais des imprimantes à marguerite *Diablo*, qui permettaient de jouer sur l'espacement des caractères au 1/120^e de pouce près.

5.2.2 Marquage de caractères

Une technique similaire — mais plus facilement détectable — consiste à marquer certains caractères d'un document. Des points peuvent par exemple être placés sous les lettres d'un texte afin de dissimuler un message. Étalées sur un texte de plusieurs pages, ces marques peuvent s'avérer relativement efficaces vis-à-vis d'un œil nonaverti. Un ordinateur n'est pas indispensable à la mise en œuvre de cette technique.

En guise d'exemple, aviez-vous remarqué le message caché dans le premier paragraphe de la section *Modulation* fine d'un texte écrit?

5.2.3 Codage sous forme d'une apparence de spam

N'importe quel texte de spam peut servir de base à de la stéganographie, sur la base d'un codage binaire simple de quasi synonymes. Par exemple pactole = 1, fortune = 0; richesse = 1, aisance = 0; succès = 1, réussite = 0; etc. Des sites du Web proposent à titre de curiosité ce genre de codage et de décodage. Des textes écrits en *langue de bois* ou en *style administratif* se prêtent particulièrement bien à l'exercice.

5.3 Message transporté dans un son

Dans les formats sonores, il existe à peu près les mêmes possibilités de cacher des messages que dans les images.

Dans un fichier sonore au format MIDI, il n'existe pas de palette de couleurs mais bien différentes pistes qui peuvent être permutées.

Dans un fichier sonore avec compression sans perte, on peut cacher de l'information dans des variations imperceptibles du son, les bits faiblement significatifs.

Dans un fichier sonore avec compression avec perte, on peut cacher de l'information dans les choix de compression.

5.4 Autres possibilités

Il est aussi possible de cacher des informations dans bien d'autres types de fichiers couramment échangés sur des réseaux telle la vidéo ou bien dans des textes (ce fut une des premières formes de la stéganographie) ou encore dans des zones d'un disque dur inutilisées par le système de fichiers.

Des informations peuvent aussi être cachées sur d'autres supports que des supports informatiques.

5.5 Exemple

6 Usage

La stéganographie est exploitable dans de nombreux domaines. Elle trouve ainsi comme application commerciale le watermarking (apposition de filigranes électroniques), technique permettant de « tatouer » un fichier électronique (pour y introduire notamment des informations utiles à la gestion des droits d'auteur).

Il ne faut pas confondre le *watermarking*, par essence invisible, avec le fait que certains formats de fichiers offrent la possibilité d'inclure des méta-informations...



Steganart

Après les attentats du 11 septembre 2001, on a soupçonné Oussama Ben Laden de transmettre ses ordres en les cachant par des procédés stéganographiques dans des images transmises ou hébergées sur internet (ces suppositions n'ont jamais été étayées par des éléments concrets). Iréf. nécessaire

Si la cryptographie, qui permet de protéger la vie privée et l'activité industrielle sans cacher cette protection, est souvent maltraitée par les États totalitaires et les sociétés démocratiques à tendance sécuritaire, il n'en va pas nécessairement de même pour la stéganographie, qui est pourtant une technique beaucoup mieux adaptée à une activité criminelle éventuelle.

7 Voir aussi

7.1 Articles connexes

- Acrostiche
- Eidesis
- Tatouage numérique
- Steghide

7.2 Liens externes

- Paru aux éditions Points, Mots d'amour secrets, 100 lettres à décoder pour amants polissons, Jacques Perry-Salkow & Frédéric Schmitter
- (en latin) Steganographia (Secret Writing) (1500) par [Johannes Trithemius] : version numérisée d'un vieux livre écrit en latin
- (en italian) Steganography from Herodotus to Bin Ladin
- OpenPuff freeware Steganographie
- VSL: Virtual Steganographic Laboratory
- Class StreamSteganography Cette utilisation libre en classe, peut être utilisée pour stocker et récupérer des informations cachées caché des images PNG dans le langage populaire PHP.

8 Notes et références

- [1] Voir frise historique, page 6.
- [2] Exemple visuel de dissimulation d'une image dans une autre en utilisant la technique des bits de poids faible
 - Portail de la cryptologie
 - Portail du renseignement
 - Portail de la sécurité de l'information

9 Sources, contributeurs et licences du texte et de l'image

9.1 Texte

• Stéganographie Source: http://fr.wikipedia.org/wiki/St%C3%A9ganographie?oldid=112448904
Contributeurs: Dtcube, Ryo, Nataraja, Fpeters, Kelson, Herman, Serged, Eon2004, TBTB, Koyuki, Symac, Roby, Cham, ZeroJanvier, Haypo, Tieno, Gotrek, Phe, MedBot, Sam Hocevar, Iznogood, Francois Trazzi, Phe-bot, François-Dominique, Rigolithe, Markadet, Tornad, Tegu, Eskimo, Dake, Bayo, Xavier Combelle, Pabix, Sador, Sherbrooke, Goldy, Ripounet, Pallas4, Korg, Madd0, Gribeco, Zetud, Vazkor, Qsinagra, FreD, David Berardan, Tuxfan79, Lmaltier, Marc.m, A3nm, Cherry, Gzen92, Solensean, Coyau, RobotQuistnix, Gpvosbot, Iunity, FlaBot, YurikBot, Eskimbot, Candelabre, Sand, MMBot, Flo, Eltrai, ObiWan Kenobi, Sylenius, Michel1961, Manu1400, Gilles MAIRET, Remsirems, WartBot, Escalabot, Thijs!bot, Grimlock, Gilles.L, Laurent Nguyen, Rémih, Éclusette, Matrix76, CommonsDelinker, RM77, Galithiel, Salebot, Bot-Schafter, Ardus Petus, TXiKiBoT, Mikayé, Chicobot, Remi.mahel, AlleborgoBot, Nameless44, SieBot, Camille Grey, Mkossa, Eunostos, Raude, Tuxlinuxien, Francis Vergne, Grizzy Kret, Alexbot, WikiCleanerBot, Maurilbert, VanBot, ZetudBot, Linedwell, Bazook, Fapp, Guillaume70, Leszek Jańczuk, MaxLanar, GrouchoBot, JmCor, Teumteum, Abracadabra, Xqbot, Robert Landon, Matei13, Jno972, EmausBot, Sisqi, IJKL, WikitanvirBot, ChuispastonBot, OrlodrimBot, Ammontami, ZipoBibrok5x10^8, Titlutin, Mattho69, Jop108, Ramzan, Addbot, Freshgod, Anonyme002, NaggoBot, Tabarinw et Anonyme: 94

9.2 Images

- Fichier:Crypto_key.png Source: http://upload.wikimedia.org/wikipedia/commons/5/5b/Crypto_key.png Licence: LGPL Contributeurs: Originally from fr.wikipedia; description page is/was here. Artiste d'origine: Original uploader was Dake at fr.wikipedia Later versions were uploaded by Croquant at fr.wikipedia.
- Fichier: Disambig_colour.svg Source: http://upload.wikimedia.org/wikipedia/commons/3/3e/Disambig_colour.svg Licence: Public domain Contributeurs: Travail personnel Artiste d'origine: Bub's
- Fichier: Eye-Brown.svg Source: http://upload.wikimedia.org/wikipedia/commons/f/f0/Eye-Brown.svg Licence: CC0 Contributeurs: dingbat fonts Artiste d'origine: unknown:simple wide-spread figure
- Fichier:Steganart_2.jpg Source : http://upload.wikimedia.org/wikipedia/commons/a/a1/Steganart_2.jpg Licence : CC-BY-SA-3.0 Contributeurs : http://www.steganart.com/
- Fichier:Steganography_original.png Source: http://upload.wikimedia.org/wikipedia/commons/a/a8/Steganography_original.png Licence: CC-BY-SA-3.0 Contributeurs: Transferred from en.wikipedia; transferred to Commons by User:Sfan00_IMG using CommonsHelper. Artiste d'origine: Original uploader was Cyp at en.wikipedia
- Fichier:Steganography_recovered.png Source: http://upload.wikimedia.org/wikipedia/commons/c/c3/Steganography_recovered.png
 Licence: CC-BY-SA-3.0 Contributeurs: Transferred from en.wikipedia; transferred to Commons by User:Sfan00_IMG using
 CommonsHelper. Artiste d'origine: Original uploader was Cyp at en.wikipedia

9.3 Licence du contenu

• Creative Commons Attribution-Share Alike 3.0