STÉGANOGRAPHIE, L'ART DE CACHER L'INFORMATION

Réalisé par : Kerkouche Lyes

Iddir Lotfi

Groupe : 1

2015/2016

Table des matières

I. Introduction

Depuis les temps anciens, les hommes communiquent entre eux, mais aussi se battent entre eux, Cela a conduit très rapidement à la formation de clans, les membres de ces derniers ayant besoin de communiquer ont développé des méthodes d'échange par écrit, mais ils avaient intérêt à ce que leur contenu ne tombe pas en main d'ennemis. Parmi ces méthodes on a vu la cryptographie, il y a aussi sa petite sœur : la stéganographie.

Comme beaucoup de sœurs, la cryptographie et la stéganographie se ressemblent, mais sont aussi fondamentalement différentes, en effet, quand la cryptographie cherche à ce que le message soit inintelligible, la stéganographie elle cherche à ce que le message lui-même ne soit jamais découvert qu'il reste secret.

II. Définition

Du grec steganos (caché) et graphos (écriture), on peut définir la stéganographie comme l'occultation d'information dans un canal couvert avec le but de prévenir la détection d'un message caché.

La stéganographie étudie la série de techniques dont le but est d'insérer l'information sensible dans un autre fichier. On appelle ce fichier le fichier conteneur (graphiques, documents, programmes exécutables, etc.). De cette façon, on obtient que l'information passe sans avertissement aux tiers, pour que soit récupérée par un utilisateur légitime qui connaisse un algorithme déterminé d'extraction de celle.

Cette science a suscité beaucoup d'intérêt au cours des dernières années du à qu'elle a été utilisée par organisations criminelles et terroristes. Néanmoins, il ne s'agit pas d'invention nouvelle, la stéganographie s'utilisait depuis l'antiquité lointaine. Cet exposé va introduire le champ de la stéganographie, éclaircissant ses différences avec la cryptographie et montrant exemples de software pour faire bon usage de cette technique.

III. Histoire et origines

Plus de 400 ans av. J.-C., Hérodote dans son livre Les Histoires a reflété l'usage de la stéganographie dans la Grèce antique. Dans le livre on décrit un personnage qui prend un cahier de deux feuilles ou planchettes; raie bien la cire qui les couvre et sur le bois grave un message et le couvre a la cire de nouveau.

Une autre histoire, dans le même livre, décrit comme un autre personnage rase avec un couteau la tête d'un de ses esclaves et lui tatoue un message sur le cuir chevelu. Puis il attend que le cheveu pousse de nouveau.

Planchette pour écrire un message caché gravé sur le bois sous la cire

Un exemple historique de plus de l'usage de la stéganographie est le livre Hypnerotomachia Poliphili de Francesco Colonna, de l'année 1499. Dans ce livre, prenant la première lettre des 38 chapitres on peut lire "Poliam frater Franciscus Columna peramavit", ce qui est traduit : "Le frère Francesco Colonna aime passionnément Polia".

Assez plus familier pour le lecteur est l'exemple de la teinte invisible. Les enfants jouant s'envoient des messages écrits avec le jus de citron ou substances similaires, de sorte que faisant chauffer la surface sur laquelle le message est écrit, le message apparait en couleur café. Cette technique peut être plus compliquée si on implique des réactions chimiques.

La stéganographie a été présente dans notre civilisation depuis les temps immémoriaux et a été traditionnellement utilisée par les agences militaires, les criminels et la police. Or, la stéganographie classique se basait uniquement sur la méconnaissance du canal couvert

utilisé, tandis que dans l'époque moderne on utilise les canaux digitaux (image, vidéo, audio, protocoles de communication, etc.) pour atteindre le but. Dans beaucoup de cas, l'objet conteneur est connu, ce qui est ignoré est l'algorithme d'insertion de l'information dans l'objet.

IV. fondements théoriques

La stéganographie est une solution du problème classique du prisonnier. Dans une prison de haute sécurité deux internés dans cellules éloignées, Romulo et Remo, veulent communiquer pour élaborer un plan d'évasion. Or, toute la communication échangée entre eux est examinée par un agent qui les isole complètement pour qu'ils n'aient aucune communication cachée. Avec la stéganographie l'agent inspecte les messages vraisemblablement inoffensifs qui contiennent un canal subliminal très utile pour les prisonniers.

On peut observer les acteurs différents impliqués dans le champ de la stéganographie :

- **Objet conteneur:** il s'agit de l'entité qui est utilisé pour porter le message caché. Retournant à l'exemple des messages sur le cuir chevelu, l'objet conteneur est l'esclave.

- **Stégo-objet:** il s'agit de l'objet conteneur et le message caché. Suivant l'exemple, il s'agit de l'esclave, on lui a écrit un message sur son cuir chevelu et le cheveu en poussant l'a caché.

- **Adversaire:** sont tous les êtres de qui on veut cacher l'information. Dans l'exemple de la prison, il s'agit de l'agent qui livre les messages aux deux prisonniers. Cet adversaire peut être passif ou actif. Un adversaire passif soupçonne qu'une communication couverte peut se produire et essaie de découvrir l'algorithme du stégo-objet, mais il n'essaie pas de modifier l'objet. Un adversaire actif, en plus d'essayer de trouver l'algorithme de communication couverte, modifie le stégo-objet avec le but de corrompre n'importe quel essaie de messagerie subliminale.

- **Stégo-analyse: science** qui étude la détection (attaques passives) et/ou annulation (attaques actives) d'information cachée en couvertures différentes, aussi la possibilité de localiser l'information utile (existence et taille).

Tenant compte de qu'il y peut avoir adversaires actifs, une bonne technique stéganographique doit être robuste face aux distorsions, accidentelles ou résultat de l'interaction d'un adversaire actif.

La robustesse face aux distorsions normalement est un objectif de la cryptographie, or, la stéganographie et la cryptographie sont des champs différents. En cryptographie, l'objectif est assurer la confidentialité de l'information face aux yeux d'un intercepteur qui est capable de voir le cryptogramme, quand il connait l'algorithme qui le génère. En revanche, la stéganographie cherche cacher la présence du message lui-même; si on arrive à identifier la position du message, on connait directement la communication, ce qui ne se passe pas dans le cas du cryptogramme.

Au début du siècle XX, Kerkhoff a formulé une série de principes qui ont été nommés comme les piliers essentiels dans le champ de la sécurité, un d'eux indique : «assume que l'utilisateur (malicieux) connait tous les procédures de chiffrage». Si on applique le principe en la stéganographie, cela signifie assumer que l'agent connait l'algorithme caché dans le message dans l'objet-conteneur, ce qui implique l'isolement immédiat des prisonniers.

Pour que la stéganographie soit plus utile il faut la combiner avec la cryptographie. Le message à échanger doit être chiffré (de façon robuste) et après introduit dans l'objet-conteneur. De cette façon, bien qu'une ennemie intercepte le message, il ne pourrait jamais connaitre le message échangé.

La combinaison de ces deux techniques a un autre avantage additionnel, quand on utilise la cryptographie en solitaire on connait que des messages s'échangent, ce qui peut servir comme un point de départ pour une attaque avec le but de découvrir le message. Quand on introduit la stéganographie, dans la majorité des cas on ne connait pas l'existence d'une communication chiffrée.

V. Fonctionnement et exemples

On se centrera sur l'objet-conteneur le plus utilisé : les images digitales. Spécialement, en format BMP pour sa simplicité (c'est un format de fichier sans compression). Les idées présentées peuvent être étendues à autres formats (JPG, PNG, etc.) et à autres conteneurs (vidéos, documents, etc.) à condition que les particularités de chaque format soient respectées.

Remplacement de bits de l'objet-conteneur

Cette technique consiste en substituer certains bits du fichier conteneur par ceux de l'information à cacher. L'avantage de cette approche est que la taille du fichier conteneur ne se voit pas altéré et grâce à la redondance et/ou excès de détaille dans les fichiers, dans beaucoup de cas pas leur qualité.

Par exemple, dans un fichier de son on peut utiliser les bits qui ne sont pas audibles par l'oreille humaine pour être remplacés par les bits du message.

Si on travaille avec les images, la méthode traditionnelle consiste en substituer les bits moins significatifs (LSB), dans une gamme de couleur de 24 bits (plus de 16 millions de couleurs). Cela se traduit qu'un pixel avec le ton rouge se voit 1% plus obscur. Dans beaucoup de cas ce sont des changements inestimables aux sens humains, et peuvent être détectés moyennant une analyse informatique de la structure des fichiers.

Les fichiers BMP sont un format standard d'image de carte de bits dans les systèmes opératifs DOS, Windows et valide pour MAC et PC. Supporte des images de 24 bits (millions de couleurs) et 8 bits (256 couleurs) et peut travailler en échelle des gris, RGB et CMYK.

Zoom illustratif sur pixel d'une image

Chaque pixel d'un fichier BMP de 24 bits est représenté par trois octets. Chacun de ces octets contient l'intensité de couleur rouge, vert et bleu (RGB : red, green, blue). Combinant les valeurs dans ces positions nous pouvons obtenir les 224, plus de 16 millions de couleurs qu'un pixel peut montrer.

À son tour, chaque octet contient une valeur entre 0 et 255, ou ce qui est le même, entre 00000000 et 11111111 en binaire, étant le chiffre de la gauche du poids supérieur. Ce qui démontre qu'on peut modifier les bits les moins significatifs d'un pixel sans produire une altération.

Effet visuel de la modification des bits les moins significatifs des composants RGB d'un pixel

Insertion de bits dans l'objet-conteneur

Dans ce cas on ajoute les bits d'information d'une marque structurelle déterminée du fichier (fin de fichier ou EOF, espaces de padding ou alignement, etc.). Cette option présente l'inconvénient de modifier la taille de l'objet-conteneur.

Pour déduire cette idée à l'exemple des images BMP il faut comprendre premièrement comment le format est structuré. Les premiers 54 octets contiennent les métadonnées de l'image, qui sont divisées de la façon suivante :

2 octets : toujours contiennent la chaîne ‘BM’, ce qui révèle qu’il s’agit d’un BMP.

4 octets : taille du fichier en octets.

4 octets : réservés (pour des utilisations futures), contiennent zéros.

4 octets : offset, distance entre le haut et le premier pixel de l’image.

4 octets : taille des métadonnées (la structure elle-même).

4 octets : largeur (numéro de pixels horizontaux).

4 octets : hauteur (numéro de pixels verticaux).

2 octets : numéro de plans de couleur.

2 octets : profondeur de couleur.

4 octets : type de compression (égale zéro, parce que BMP est un format pas comprimé).

4 octets : taille de la structure image.

4 octets : pixels par mètre horizontal.

4 octets : pixels par mètre vertical.

4 octets : quantité de couleurs utilisés.

4 octets : quantité de couleurs importants.

Avec cette structure, la forme triviale de cacher les données consiste en les cachant après les métadonnées (entre les métadonnées et les données de l’image) et modifier le champ offset (distance entre les métadonnées et les pixels de l’image). De cette façon, on peut laisser un espace pour tout le contenu additionnel qu’on veut accueillir.

VI. Stégo-analyse

Comme on a mentionné déjà, la stégo-analyse est la technique utilisée pour récupérer les messages cachés ou pour empêcher la communication par stéganographie. Il y a deux types principaux de stégo-analyse passive :

Stégo-analyse manuelle

Consiste à chercher de façon manuelle des différences entre l’objet conteneur et le stégo-objet cherchant des changements dans la structure pour localiser les données cachées. Les principales inconvenances de cette technique sont qu’il est nécessaire d’avoir l’objet conteneur et que dans beaucoup de cas on détecte qu’un objet contient information cachée mais il est impossible de la récupérer. Néanmoins, quand on n’a pas le fichier conteneur, on

peut chercher les irrégularités dans le fichier stéganographié pour essayer de trouver les signes de l'existence des données cachées.

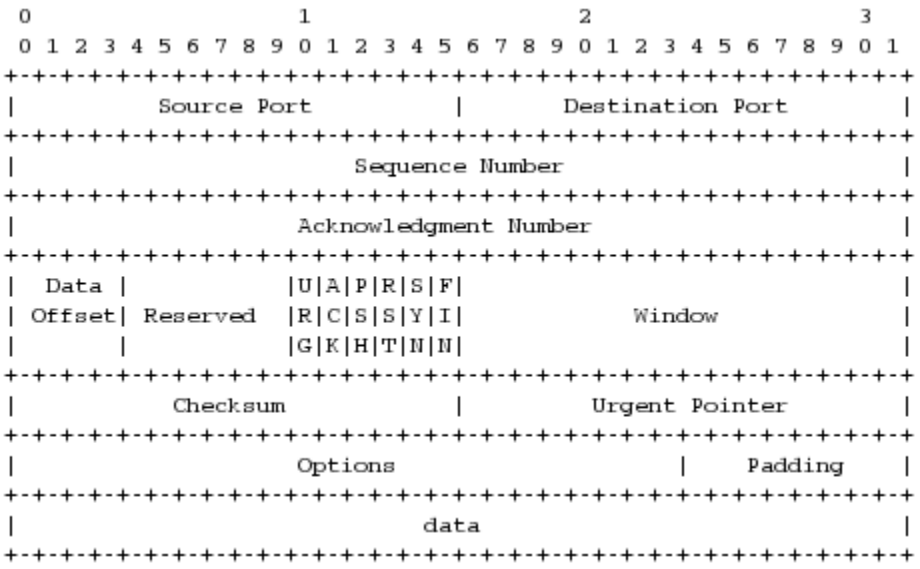
Les attaques visuelles avertissent l'oeil humain de la présence d'information cachée grâce à l'application de filtres. Considérez le cas du BMP où le bit le moins significatif des composants de quelques pixels a été substitué par l'information cachée. Dans ce contexte la stégo-analyse manuelle consiste en appliquer un filtre de sorte qu'on considère seulement le bit le moins significatif de chaque composant RGB de chaque pixel.

Stégo-analyse statistique

Consiste en la confrontation de la fréquence de distribution de couleurs du stégo-objet. C'est une technique lente pour laquelle il faut utiliser software spécialisé. Ces programmes normalement cherchent des règles pour occulter les messages utilisés par les programmes les plus habituels de stéganographie, cette approche les fait très efficients quand il s'agit de messages cachés avec ces programmes typiques. Or, les messages cachés manuellement sont presque impossibles de trouver pour ces programmes.

VII. Applications de la stéganographie

Stéganographie utilisant le protocole TCP/IP



Haut du protocole TCP

Par exemple, considérant uniquement le haut TCP, on peut occulter les données dans le numéro de séquence initiale d'une connexion. Cela offert 32 bits de données cachées par paquet de connexion initiale (SYN), c'est-à-dire, 4 caractères ASCII. Suivant cette philosophie on peut occulter information dans autres champs des hauts des protocoles différents qui composent TCP/IP, à condition que les changements n'impliquent pas le refus des paquets échangés.

Tatouages numériques

Un tatouage numérique est un code d'identification qui est introduit directement dans le contenu d'un fichier multimédia, normalement pour inclure information relative aux droits d'auteur ou de propriété du contenu numérique en question.

La présence de ce tatouage doit être inestimable pour le système de perception humaine en même temps que facilement amovible par une application télématique qui connait l'algorithme pour la récupérer.

Les applications les plus communes des tatouages sont :

Preuve de propriété: identification de la source, l'auteur, le propriétaire, le distributeur et/ou le consommateur d'un fichier numérique.

Empreinte digitale (*fingerprinting*): inclut les données liées à une transaction, données du propriétaire d'un fichier et de son acheteur. Permet d'identifier le responsable des copies illégales de contenu protégé par droits d'auteur.

Classification de contenus: les tatouages peuvent être utilisés pour indiquer le type de contenu d'un fichier.

Autres applications

Autre utilisation peu éthique de la stéganographie est la fuite d'information dans les environnements organisationnels, militaires, gouvernementaux, etc. Dans endroits où on inspecte le contenu qu'un employé extrait de l'environnement en moyens numériques, la stéganographie peut être utilisée pour porter schèmes, documents et autre information délicate.

VIII. Conclusions

La stéganographie est une technique en évolution constante, avec une historie longue et avec la capacité de s'adapter aux nouvelles technologies. À mesure que les outils de stéganographie deviennent plus avancés, les techniques et les outils utilisés en stégo-analyse aussi deviennent plus complexes.

Les fichiers conteneurs ne doivent pas être forcément images, n'importe quel moyen est valide (audio, vidéo, exécutables, etc...)

Les nouvelles techniques de stégo-analyse font usage de la stéganographie combinée avec la cryptographie avec le but d'atteindre un niveau de sécurité raisonnable. La cryptographie garantit la confidentialité d'une conversation mais ne cache pas le fait que la conversation est maintenue. D'autre côté, la stéganographie en solitaire peut cacher le fait qu'une conversation est maintenue, mais quand l'interaction est découverte, est possible qu'un ennemie connaisse le contenu échangé. Bien qu'il soit difficile de découvrir le contenu original, une ennemie peut modifier le stégo-objet pour empêcher la communication.

IX. Webographie

http://www.bibmath.net/crypto/index.php?action=affiche&quoi=stegano/tatouage

https://www.securiteinfo.com/attaques/divers/steganographie.shtml

https://fr.wikipedia.org/wiki/St%C3%A9ganographie

https://fr.wikipedia.org/wiki/Tatouage_num%C3%A9rique

http://www.univ-orleans.fr/mapmo/membres/louchet/teaching/timo/Rogerie.pdf

http://www.bibmath.net/crypto/index.php?action=affiche&quoi=stegano/histstegano