Compte rendu INFO002 - Cryptologie
Le tatouage numérique

Juliette BLANC
Thomas CARRERA

<u>Sommaire</u>

Introduction

Ce document rend compte des recherches effectuées dans le cadre du cours de Cryptologie INFO002. Pour rappel, le sujet choisi est le tatouage numérique.

Utile aux organisations pour déterminer la source d'une fuite de données, qu'elle soit volontaire ou du fait d'une négligence, le tatouage de données numériques permet de tracer la provenance d'informations ou de fichiers. Mais plus exactement, qu'est-ce que le tatouage de données ? Quels sont ses principes de fonctionnement ? Et de quelles manières est-il possible de l'utiliser ?

I. Qu'est-ce que le tatouage numérique ?

A. Quelques définitions

Tatouage numérique : Le tatouage numérique est une technique permettant d'ajouter des informations ou des messages à un fichier, un signal (audio, vidéo ou image) ou un autre fichier numérique. Le plus souvent, ces informations sont en lien avec le fichier en question et sont également imperceptibles.

Stéganographie : La stéganographie est l'art de dissimuler au sein d'un support anodin une information qui, bien souvent, est sans rapport avec le support. Cette dissimulation se fait de sorte qu'il soit difficile pour un observateur extérieur de se rendre compte qu'il y a eu dissimulation.

Cryptologie : La cryptographie est l'art de rendre indéchiffrable un message et ceci au sus de tout le monde.

La définition du tatouage le rapproche beaucoup de la stéganographie que de la cryptologie. Cependant, il faut garder à l'esprit les différences essentielles entre ces deux techniques.

En particulier, il est indispensable en stéganographie que l'existence même du message soit dissimulée, alors qu' au contraire, en tatouage, la connaissance publique de l'existence d'une marque dans un document peut être un moyen de dissuasion contre le piratage.

Par ailleurs, en stéganographie, le message de couverture n'est pas important en soi, alors qu'en tatouage, il est primordial que ce message ne soit pas dénaturé, à la fois lors de l'insertion du tatouage et lors d'une attaque visant à détruire la marque. Le lien entre le message caché et les données support est donc beaucoup plus fort.

B. L'histoire du tatouage numérique

Le tatouage numérique peut se rapprocher de la stéganographie dans son fonctionnement. La stéganographie est apparue il y a très longtemps, vers 445 avant J.-C. Quant au tatouage numérique, il faudra attendre les années 1990 pour que son utilisation soit plus répandue.

Les premiers travaux sur cette technique étaient motivés par les problèmes de défense du droit d'auteur (copyright) dans un environnement numérique ouvert. La duplication sans perte de qualité et la rapidité de diffusion dans un environnement tel internet faisaient que toute œuvre numérique (image, film, musique, logiciel…) pouvait être copiée et distribuée extrêmement facilement sans contrôle des ayants-droit.

C. Les différents types de tatouage numérique

Il est possible de distinguer le tatouage numérique suivant deux types différents : le tatouage visible et l'invisible.

Les techniques de tatouage visible altèrent le signal ou le fichier (par exemple ajout d'une image pour en marquer une autre). Il est, par exemple, très souvent utilisé par les agences de photo. Ces dernières vont ajouter un tatouage visible en forme de copyright ("©") aux versions de prévisualisation (basse résolution) de leurs photos, dans le but d'éviter que ces versions ne se substituent aux versions hautes résolutions payantes.

Les techniques de tatouage invisible, quant à elles, modifient le signal d'une manière imperceptible par l'utilisateur final (par exemple en ne modifiant que le bit le moins significatif de chaque octet).
Si nous reprenons l'exemple des agences de photo, les photos hautes résolutions vendues par l'agence peuvent posséder un tatouage invisible, qui ne dégrade pas le contenu visuel. Ce tatouage pourra alors permettre de détecter l'éventuelle source d'un vol.

Enfin, il existe aussi des tatouages dits fragiles. Ce sont des tatouages invisibles, qui sont utilisés pour détecter toute modification du signal, dans le but, par exemple, de vérifier que le contenu n'a pas été modifié par un tiers.
Certains types de tatouages fragiles peuvent également servir à détecter des contrefaçons de documents physiques. Pour cela, on peut intégrer ces tatouages dans des images qui sont ensuite imprimées.

Dans la suite de ce compte-rendu, nous nous baserons principalement sur les tatouages invisibles.

D. Les différents niveaux de tatouage numérique

Voici une liste non exhaustive des différents niveaux de tatouage numérique :

Tatouage fragile : Le tatouage en question ne résiste à rien sert surtout pour détecter des modifications de signal.

Tatouage robuste : Le tatouage résiste à la transformation non malveillante (compression).

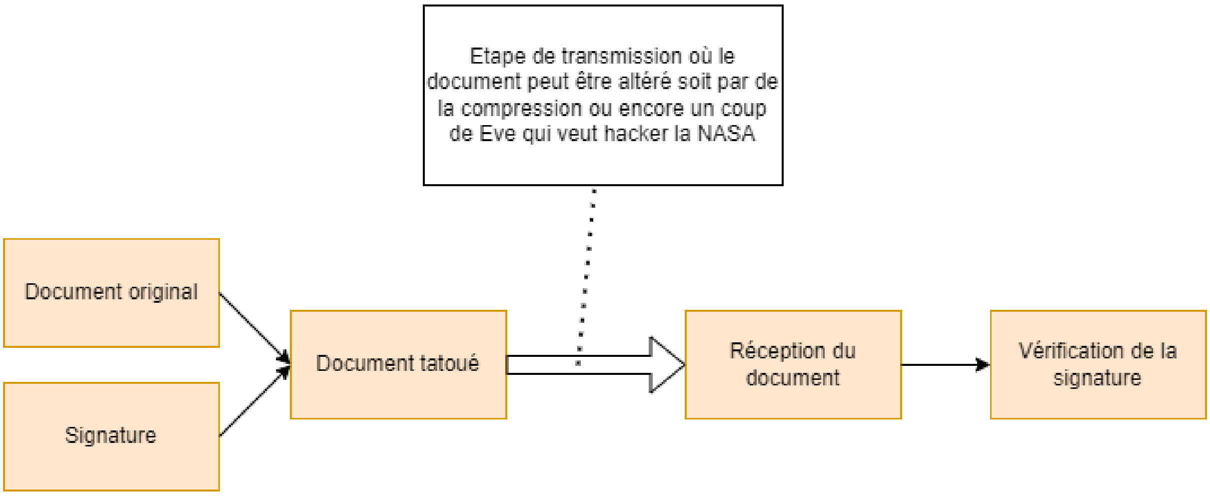
Tatouage sécurisé : Le tatouage résiste à la transformation malveillante.

Tatouage enrichissant : Le tatouage contient des données pour enrichir le document.

Tatouage réversible : Il s'agit d'un tatouage fragile qui permet de retrouver l'image originale.

II. Comment fonctionne le tatouage numérique ?

Voici le grandes étapes dans le processus du tatouage numérique :



La première étape consiste à insérer la signature dans le document (il y a toute une partie mathématique que nous n'allons pas aborder ici).

Une fois le document tatoué, il va être "lâché dans la nature" comme par exemple mis sur Facebook et va subir tout un tas de transformation comme de la compression (qui sur certaine plateforme peut être très importante). Il pourra également être victime de personnes malveillantes qui pourraient vouloir usurper un copyright par exemple.

La dernière étape après la réception du document va être de vérifier la signature pour connaître l'authenticité du document. Selon les transformations subies, la vérification peut ne pas être toujours juste.

III. Comment peut-il être utilisé ?

A. Identification du propriétaire

Le tatouage numérique permet d'identifier l'ayant droit du support à la manière du copyright qui est une forme de tatouage numérique visible. Cependant cette méthode présente des différences avec un tatouage numérique invisible.

En effet, le tatouage numérique est invisible donc moins voyant que le copyright, ce qui le rend également plus sûr.

Cela a tout de même des désavantages car du fait de son invisibilité cela n'a aucune valeur devant une cours de justice et il sera également plus difficile de contacter l'ayant droit de l'œuvre pour pouvoir l'utiliser à son tour.

Enfin, il se peut parfois que les systèmes n'extraient pas correctement le tatouage notamment à cause de dégradations physiques par exemple.

B. <u>Preuve de propriété</u>

Nous avons vu précédemment qu'il était facile de marquer un document cependant il est beaucoup plus difficile de prouver qu'il nous appartient. En effet, un attaquant peut remplacer le copyright et dire que le document lui appartient.

Il y a deux solutions à ce problème :

- Il faut enregistrer son copyright à une office de copyrights, ce qui coûte assez cher surtout s' il y a beaucoup de documents à enregistrer.
- Il faut montrer que l'on détient l'original en démontrant qu'un des deux documents et un dérivé de l'autre. Ce qui n'est pas forcément facile.

C. <u>Authentification du support</u>

La présence du tatouage numérique permet de savoir si le message reçu est altéré ou non (une signature en cryptologie), l'avantage est que la signature se trouve dans le document donc pas de risque de perte de la signature.

Selon l'utilisation on peut utiliser plusieurs types de tatouage :

- Le tatouage fragile : on peut signer localement le document (par blocs) ce qui permet de détecter la zone dégradée.
- Le tatouage robuste : en le rendant plus robuste, le tatouage permet de résister à la compression avec perte qui peut être utile dans le cas d'une image.

D. La société Lamark

Lamark à créé une plateforme nommée Imatag https://www.imatag.com/ qui propose un service permettant de marquer et de détecter les œuvres pour les photographes.
La plateforme va scanner des millions de sites web à la recherche d'œuvres et va les indexer. Ainsi les photographes pourront savoir qui utilise leurs œuvres et savoir si leurs utilisations sont licites.

La société a d'ailleurs été sélectionnée par le ministère de la culture pour créer un moteur de recherche d'images.

E. Le domaine de l'intelligence artificielle

eIQ Model Watermarking est un outil créé par la société NXP qui permet de tatouer un modèle IA pour ajouter le copyright de protection de droit d'auteur. Cela permet aux développeurs de prouver qu'un modèle d'apprentissage automatique est en réalité une copie ou un clone du logiciel représentatif de leur propriété intellectuelle, et ce sans qu'ils aient besoin d'accéder au code du modèle incriminé.

C'est une solution intéressante car aujourd'hui peu de protections contre le droit d'auteur sont mises en place dans le domaine de l'intelligence artificielle.