Stéganographie dans les images

Victoria D'AURA
Damien TOOMEY

INSA – Institut National des Sciences Appliquées de Rouen

10 décembre 2018

Sommaire

Sommaire

Introduction

Stéganographie – Etymologie

du grec *steganos*, caché ; et *graphein*, écrire

Stéganographie – Définition

Technique qui consiste à cacher une information en la dissimulant au sein d'une autre information.

Différents types de stéganographie

- dans des textes
- dans des fichiers audio
- dans des fichiers vidéo
- dans des fichiers systèmes
- **dans des images**

Sommaire

Origines et Applications (1/2)

Origines

La stéganographie existe depuis longtemps, bien avant l'invention de l'ordinateur. (500 av. J.C.)

Mode de propagation des malwares

Ransomware SyncCrypt (août 2017)
Une image contenant un .zip se décompresse à l'ouverture de l'image puis lance un exécutable pour chiffrer les données de l'ordinateur cible.

Origines et Applications (2/2)

Mode de signature : le Watermarking

- Cacher un copyright au sein d'une oeuvre protégée
- Prouver l'authenticité d'un billet de banque
- Prouver l'intégrité d'un fichier

Limites de la stéganographie

L'information dissimulée dans le document est dépendante de la nature de ce document.

Exemple

Une information dissimulée dans une image PNG (format sans perte d'information) peut être détruite si l'image est convertie en JPG (format avec perte d'information).

Sommaire

Images RVB

Acronyme

- R : Rouge
- V : Vert
- B : Bleu

Définition

$$\textit{Image RVB} = \textit{empilement}(\textit{Matrice Rouge}, \textit{Matrice Verte}, \textit{Matrice Bleue})$$

$$1\ \textit{pixel} = 3 \cdot 8\ \textit{bits} = 24\ \textit{bits} \quad \text{soit } 8\ \textit{bits} \text{ par couleur}$$

Chaque couleur (R,V et B) a une palette d'entiers entre 0 et 255, ce qui permet d'avoir des nuances d'une même couleur.

Sommaire

Principales méthodes de stéganographie dans les images

Stéganographie dans le domaine spatial

1. LSB : Least Significant Bit
2. ELSB : Enhanced/Edge Least Significant Bit Embedding
   (aussi appelé Masking and Filtering)
3. RLSB : Random Least Significant Bit Insertion

Stéganographie dans le domaine fréquentiel

Méthode générale

LSB : Least Significant Bit (1/3)

Description théorique

Principe : Codage/Décodage

1. Fonction de position : obtenir la position du pixel qu'on va modifier
2. Insertion/Extraction du message aux positions données par la fonction de position

Remarques

Cette méthode est vulnérable à la stéganalyse.
Dans la réalité, le message est chiffré avant de le cacher dans l'image

LSB : Least Significant Bit (2/3)

Explication de l'implémentation - **Fonction de position**

Soit le message : "Il fait beau aujourd'hui" à cacher dans l'image.

$$\text{1ère lettre du message} = \text{I}; \text{codeAscii(I)} = 73; \text{codeBinaire}(73) = \underbrace{01001001}_{8bits}$$

$$\text{Fonction de position} : \begin{cases} f(x) = 2 \cdot x + 1 \\ f(y) = 2 \cdot y + 1 \end{cases}$$

$$\text{Itération 1 (1er caractère du message)} \Rightarrow \begin{cases} f(1) = 3 \\ f(1) = 3 \end{cases}$$

Méthode

Pour chaque caractère du message, on met chaque bit de ce caractère dans le bit de poids faible de chaque couleur des trois pixels consécutifs donnés par la fonction de position

LSB : Least Significant Bit (3/3)

Explication de l'implémentation - **Insertion du message**

$$\text{codeBinaire}(I) = \underbrace{01001001}_{8bits}$$

$$\begin{pmatrix} 92 & 37 & 41 & 102 \\ 37 & 37 & 40 & 31 \\ 104 & 33 & 111 & 106 \\ 108 & 34 & 113 & 32 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 01011100 & 00100101 & 00101001 & 01100110 \\ 00100101 & 00100101 & 00101000 & 00011111 \\ 01101000 & 00100001 & 0110111\,0 & 01101010 \\ 01101100 & 00100010 & 01110001 & 00100000 \end{pmatrix}$$

$$\begin{pmatrix} 32 & 37 & 41 & 102 \\ 37 & 34 & 40 & 104 \\ 31 & 33 & 37 & 106 \\ 108 & 111 & 113 & 92 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 00100000 & 00100101 & 00101001 & 01100110 \\ 00100101 & 00100010 & 00101000 & 01101000 \\ 00011111 & 00100001 & 0010010\,1 & 01101010 \\ 01101100 & 01101111 & 01110001 & 01011100 \end{pmatrix}$$

$$\begin{pmatrix} 108 & 37 & 41 & 102 \\ 37 & 37 & 40 & 104 \\ 31 & 33 & 34 & 106 \\ 32 & 111 & 113 & 92 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 01101100 & 00100101 & 00101001 & 01100110 \\ 00100101 & 00100101 & 00101000 & 01101000 \\ 00011111 & 00100001 & 0010001\,0 & 01101010 \\ 00100000 & 01101111 & 01110001 & 01011100 \end{pmatrix}$$

ELSB : Edge Least Significant Bit Embedding (1/4)

Description théorique

Principe : Codage/Décodage

1. Masquage : on met les 2 bits de poids faible du rouge à 0 pour chaque pixel
2. Détection des contours dans l'image masquée
3. Insertion/Extraction du message dans l'image originale aux coordonnées des contours de l'image masquée

Avantage : Version améliorée de l'algorithme LSB

- moins vulnérable à la stéganalyse que le LSB

Inconvénient

- algorithmes de codage/décodage sont plus longs

ELSB : Edge Least Significant Bit Embedding (2/4)

Explication de l'implémentation - **Masquage**

Matrice Rouge

$$\begin{pmatrix} 92 & 37 & 41 & 102 \\ 37 & 37 & 40 & 31 \\ 104 & 33 & 111 & 106 \\ 108 & 34 & 113 & 32 \end{pmatrix}$$

Matrice Rouge en binaire

$$\Leftrightarrow \begin{pmatrix} 01011100 & 00100101 & 00101001 & 01100110 \\ 00100101 & 00100101 & 00101000 & 00011111 \\ 01101000 & 00100001 & 01101111 & 01101010 \\ 01101100 & 00100010 & 01110001 & 00100000 \end{pmatrix}$$

Matrice Rouge en binaire après masquage

$$\Rightarrow \begin{pmatrix} 01011100 & 00100100 & 00101000 & 01100100 \\ 00100100 & 00100100 & 00101000 & 00011100 \\ 01101000 & 00100000 & 01101100 & 01101000 \\ 01101100 & 00100000 & 01110000 & 00100000 \end{pmatrix}$$

ELSB : Edge Least Significant Bit Embedding (3/4)

Explication de l'implémentation - **Détection de contours**

Image Originale

Contours de l'image masquée

ELSB : Edge Least Significant Bit Embedding (4/4)

Explication de l'implémentation - **Insertion du message dans les contours**

Choix dans l'implémentation

- on ne modifie que les deux bits de poids faible du rouge
- on cache le message dans l'image originale aux coordonnées des contours de l'image masquée

Choix dans l'implémentation

Soit le message : "Il fait beau aujourd'hui" à cacher dans l'image.

$$\text{1ère lettre du message} = \text{I}; \text{codeAscii(I)} = 73; \text{codeBinaire}(73) = \underbrace{01001001}_{8bits}$$

$\Rightarrow$ 4 pixels sont nécessaires pour cacher chaque caractère du message

Stéganographie dans le domaine fréquentiel

Description théorique

Du Spatial au Fréquentiel

Inconvénients des méthodes relatives au domaine spatial : dégradation de l'image, particulièrement lors d'une compression.

En Pratique

1. Passage du domaine spatial au domaine fréquentiel :
   - Transformée en Cosinus Discrète (TCD)
   - Transformée de Fourier Discrète (TFD)
2. Le message est intégré aux coefficients transformés
3. Les données sont repassées en spatial (TCD inverse ou TFD inverse)

La TCD bidimensionnelle est la transformée la plus utilisée dans le traitement d'images. Une méthode l'exploitant est celle dite "d'incorporation Jpeg-Jsteg", où le message est intégré aux LSB des coefficients TCD dont les valeurs sont différentes de 0, 1 et -1.

Sommaire

Stéganalyse : détection de la stéganographie

Définition

Vérifier si un document contient une information cachée.

Plusieurs types de steganalyses

- Passive : Seulement observer le trafic entre l'expéditeur et le destinataire.
- Active : Apporter des modifications à l'image (Compression, filtrage...) avec pour but de détruire le processus stéganographique s'il existe.
- "Malicieuse" : Comprendre la technique stéganographique, extraire le message, et l'utiliser à ses propres fins.

Autres types de steganalyses

- Analyser le document pour des méthodes de stéganographie connues (LSB, ELSB, RLSB,...)
- Si on possède une version de l'image originale : comparer cette image avec l'image qui contient potentiellement de la stéganographie

Sommaire

Sommaire

Conclusion

On peut cacher différents types d'information dans une image : texte, audio, image, fichier,...

Différence entre la stéganographie et le tatouage d'image

- Stéganographie : cacher une information dans un document dans le but qu'elle soit indétectable
- Tatouage d'image : cacher une information dans un document dans le but qu'elle soit ni retirable, ni modifiable

La stéganographie aujourd'hui

Chiffrage de l'information avant de la dissimuler dans une autre

Question

Pourquoi cacher le message sur le bit ou les 2 bits de poids faible de l'image ?

Sommaire

Bibliographie (1/3)

- Analysis and Implementation of Distinct Steganographic Methods

  https://arxiv.org/pdf/1108.2153.pdf
- Dissimulation de données : La stéganographie

  https://www.securiteinfo.com/attaques/divers/steganographie.shtml
- Ondelette : Tatouage et Stéga-analyse

  https://jnsecurite2018.sciencesconf.org/data/pages/slides_philippe_carre.pdf
- La stéganographie au cours des siècles

  http://www.bibmath.net/crypto/index.php?action=affiche&quoi=stegano/histstegano
- Some New Methodologies for Image Hiding using Steganographic Techniques

  https://arxiv.org/pdf/1211.0377.pdf

Bibliographie (2/3)

- A Survey on LSB Based Steganography Methods
  https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiEsdeQg_reAhWk34UKHShRAj8QFjAAegQICRAC&url=https%3A%2F%2Fwww.ijecs.in%2Findex.php%2Fijecs%2Farticle%2Fdownload%2F1912%2F1767%2F&usg=AOvVaw3GtnKqK9NY-hwX8vY-i69l
- EDGE Based Image Steganography for Data Hiding
  https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=2ahUKEwiB6qrwl_reAhUOYxoKHRF-DmcQFjACegQIBBAC&url=https%3A%2F%2Fpen2print.org%2Findex.php%2Fijr%2Farticle%2Fdownload%2F9461%2F9127&usg=AOvVaw1tXVWLzjun-sY6lRSgbllH
- Steganography and Watermarking - FI MUNI
  https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=20&ved=2ahUKEwj179bM_IvfAhUMQRoKHU6MDg4QFjATegQICBAC&url=https%3A%2F%2Fwww.fi.muni.cz%2Fusr%2Fgruska%2Fcrypto09%2FCRYPTO0911a.ppt&

Bibliographie (3/3)

- A Secure Steganographic Algorithm Based on Frequency Domain for the Transmission of Hidden Information

  https://www.hindawi.com/journals/scn/2017/5397082/

- Le marquage dans le domaine fréquentiel

  http://cours-info.iut-bm.univ-fcomte.fr/wiki/pmwiki.php/Imagerie/Steganographie#toc10

- Dissimulation de Données

  https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwjFnIrs5JLfAhVMRBoKHRnEAH4QFjAAegQIABAC&url=https%3A%2F%2Fwww.lirmm.fr%2F~wpuech%2Fenseignement%2Fmaster_informatique%2FCompression_Insertion%2FDissimulation_de_donnees_Cours3.pdf&usg=AOvVaw02tZo9hxXDCYPS8pWcdRfy