UNIVERSITÉ MONTPELLIER 2
SCIENCES ET TECHNIQUES

La stéganographie :

Introduction générale
Propriétés d'un schéma stéganographique

Méthodes d'insertion usuelles:

Insertion dans le domaine spatial
Insertion dans le domaine transformé

Méthodes d'insertion adaptatives:

Le principe des algorithmes adaptatifs
L'algorithme HUGO
L'approche treillis (STC)

La stéganalyse :

Définition de la stéganalyse
Principaux scénarios en stéganalyse
La stéganalyse ciblée
La stéganalyse aveugle
La stéganalyse sous d'autres angles

CONTACT :
Email : kouider@lirmm.fr
http://www2.lirmm.fr/~kouider

Stéganographie et stéganalyse

UNIVERSITÉ MONTPELLIER 2
SCIENCES ET TECHNIQUES

La stéganographie :

Introduction générale
Propriétés d'un schéma stéganographique

Méthodes d'insertion usuelles:

Insertion dans le domaine spatial
Insertion dans le domaine transformé

Méthodes d'insertion adaptatives:

Le principe des algorithmes adaptatifs
L'algorithme HUGO
L'approche treillis (STC)

La stéganalyse :

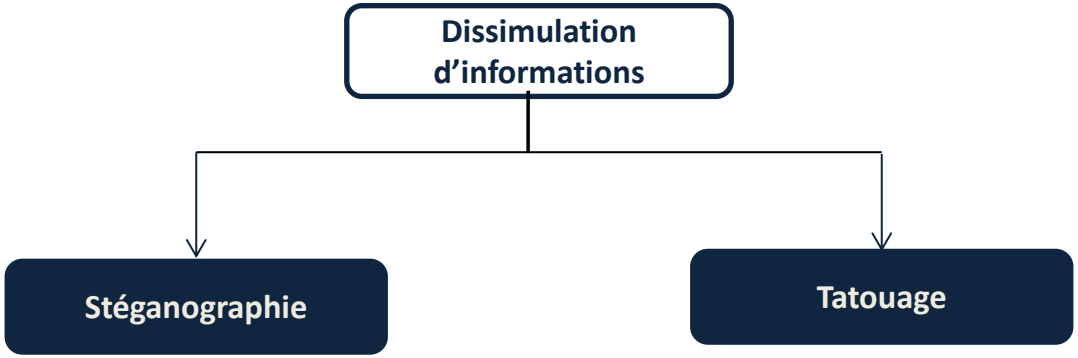
Définition de la stéganalyse
Principaux scénarios en stéganalyse
La stéganalyse ciblée
La stéganalyse aveugle
La stéganalyse sous d'autres angles

CONTACT :
Email : kouider@lirmm.fr
http://www2.lirmm.fr/~kouider

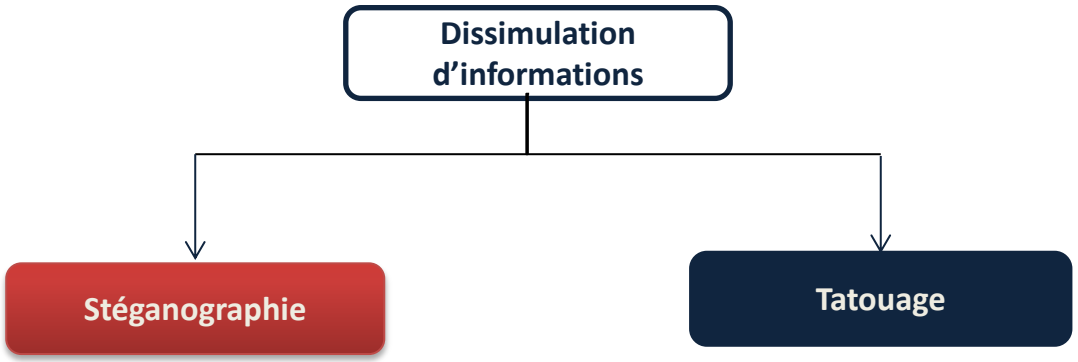
Stéganographie et stéganalyse

La stéganographie l'art de communication secrète

Techniques de dissimulation de données secrètes



Techniques de dissimulation de données secrètes



La stéganographie (Steganography) - Définition

La stéganographie

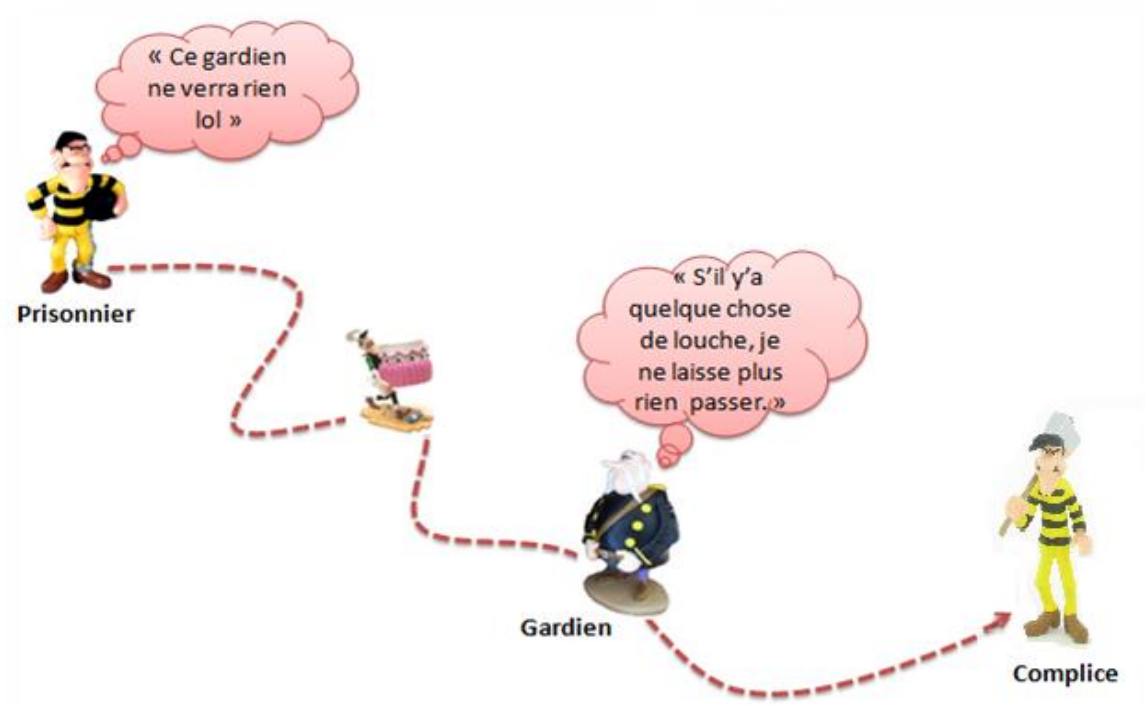
La stéganographie est l'art de communication secrète. L'objectif est de dissimuler un message secret dans **un médium anodin** (une image, une vidéo, un son..) **de sorte qu'il ne puisse être détecté (visuellement** mais aussi **statistiquement)**.

Image Hôte

≈

Image Stego

Stéganographie vs Stéganalyse

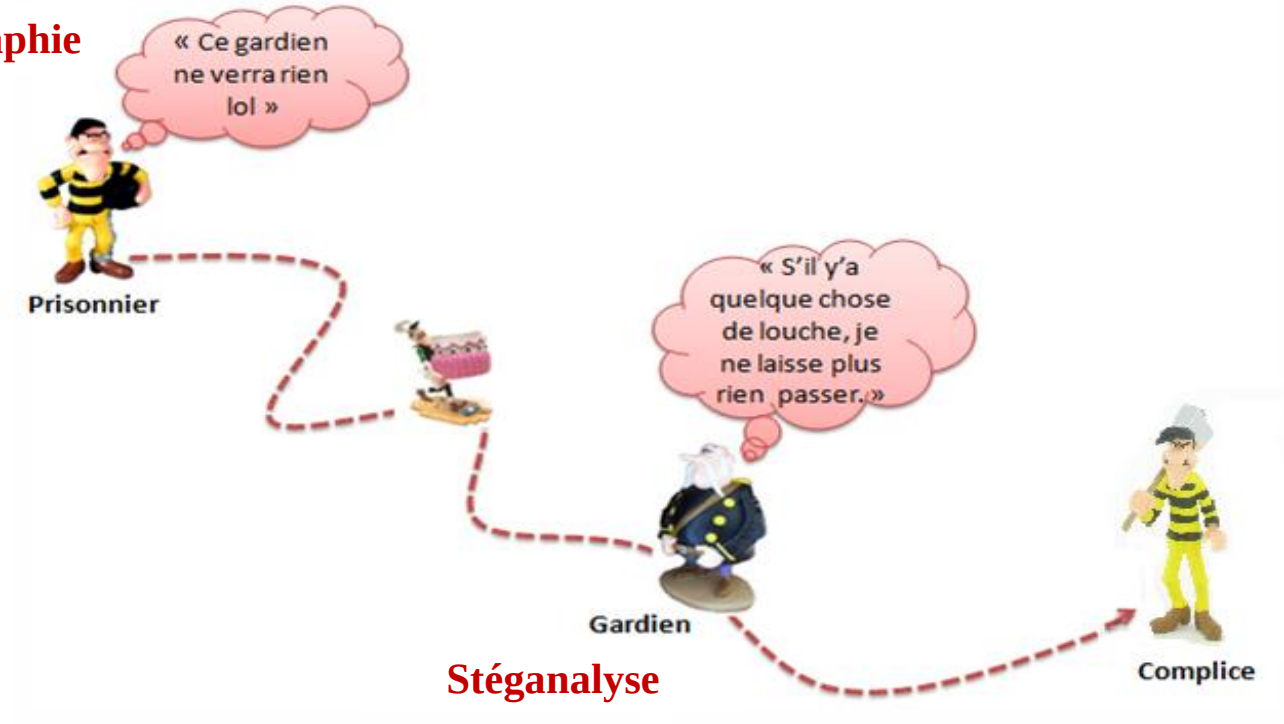


G. J. Simmons

The prisoner's problem and the subliminal channel. *In Advances in Cryptography*, CRYPTO'83, pages 51-67, Santa Barbara, CA, August 22-24,1983.

Stéganographie vs Stéganalyse

Stéganographie



Stéganalyse

G. J. Simmons

The prisoner's problem and the subliminal channel. *In Advances in Cryptography*, CRYPTO'83, pages 51-67, Santa Barbara, CA, August 22-24,1983.

La stéganographie au cours des siècles – Quelques anecdotes

- **Le tatouage de crâne des esclaves pour la communication des messages secrets.**
- **Utilisation de l'encre sympathique pour la stéganographie.**
- **La stéganographie linguistique (exemple l'acrostiche, …).**
- **L'apparition de la technique du micropoint de Zapp durant la Seconde Guerre mondiale.**
- **…**

La stéganographie de nos jours (dite moderne ou numérique)

- Avènement de l'Internet dans les années 90 et début de l'air de la stéganographie numérique.
- Utilisation des fichiers numériques (textes, vidéos, images, sons, …) pour la stéganographie.

Quelques dates

- 1996 : première conférence en stéganographie (naissance de la première communauté stéganographique).
- 1999-2001 : Premiers algorithmes scientifiques :pour JPEG : Outguess, JPHide&JPSeek, F5 ..
- 2009 : Codage matriciel (utilisation de codes correcteurs)basé BCH : [Zhang et al. 2009], [Sachnev et al. 2009],basé Reed-Solomon : [Fontaine and Galand 2009].

Stéganographie vs Tatouage vs Cryptographie

La 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 hôte**. Cette dissimulation se fait de sorte que la présence même du message soit insoupçonnée. Autrement dit, la dissimulation doit être **indétectable visuellement et statistiquement**.

Le tatouage numérique

Le tatouage est l'art d'altérer un média (un texte, une image, un son, une vidéo...) de sorte qu'il contienne un message le plus souvent **en rapport avec le média** et le plus souvent de manière **imperceptible** et **robuste**.

La Cryptographie

La cryptographie est l'art de rendre **indéchiffrable** un message et ceci au sus de toute personne tierce.

Philosophies de conception d'un schéma stéganographique

- **Stéganographie par sélection du médium de couverture.**
- **Stéganographie par synthèse du médium de couverture.**
- **Stéganographie par modification du médium de couverture.**

Philosophies de conception d'un schéma stéganographique

- **Stéganographie par sélection du médium de couverture.**
  - \+ Méthodes quasiment indétectables.
  - \+ Facile à implémenter avec une complexité faible.
  - \- Capacité d'insertion très limitée.

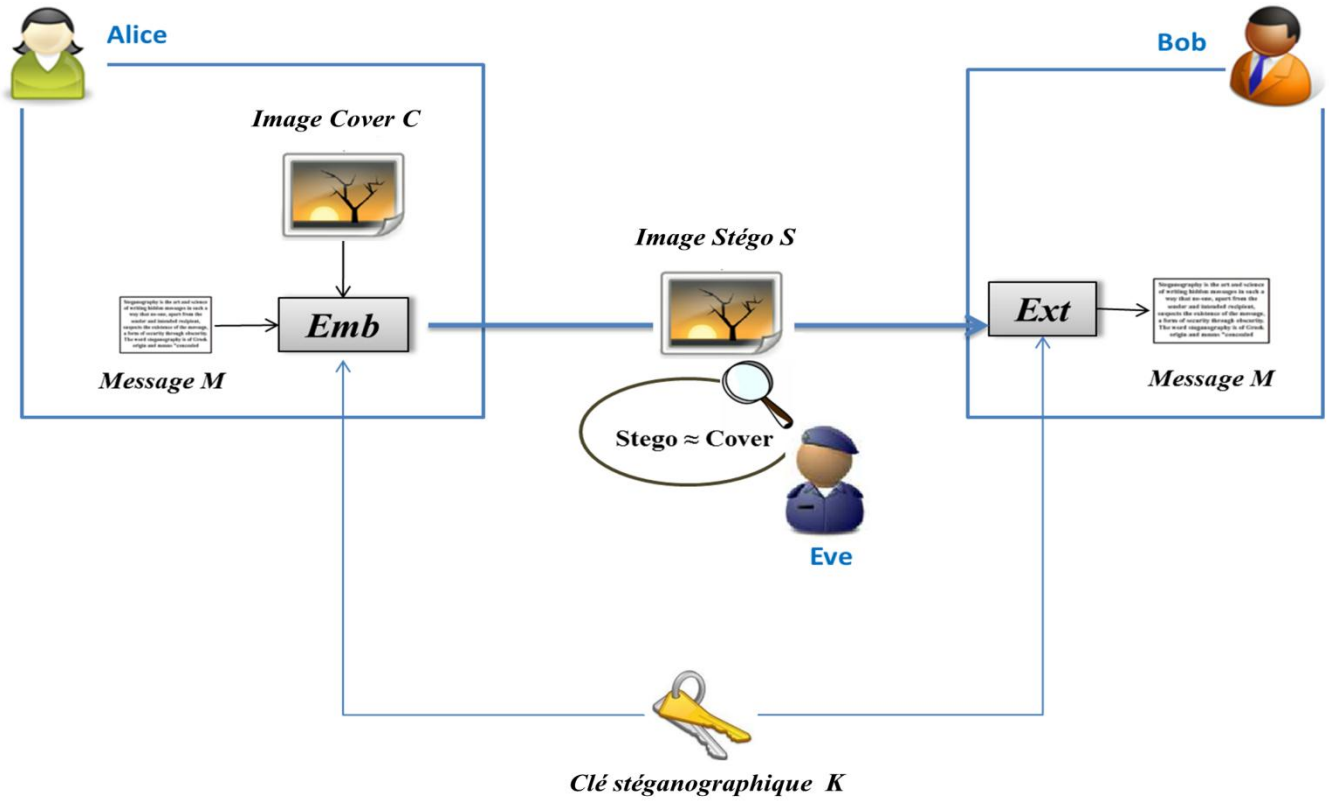
Philosophies de conception d'un schéma stéganographique

- **Stéganographie par sélection du médium de couverture.**
- **Stéganographie par synthèse du médium de couverture.**
  - \+ Méthodes parfaitement sûre.
  - \- Méthodes très théoriques, et complexes à mettre en oeuvre.

Philosophies de conception d'un schéma stéganographique

- **Stéganographie par sélection du médium de couverture.**
- **Stéganographie par synthèse du médium de couverture.**
- **Stéganographie par modification du médium de couverture.**
  - \+ Méthodes efficaces.
  - \- Plus pratiques à mettre en place.

La stéganographie par modification du médium de couverture



Étapes de dissimulation d'un message secret pour un schéma stéganographique par modification

Propriétés d'un schéma stéganographique

- **La sécurité d'un schéma stéganographique**

Point de vue théorique (Cachin 1998) :

Comparaison de la distribution du support avant et après l'insertion en utilisant la KL divergence.

$$D_{KL}(P_C \| P_S) = \sum_{x \in C} P_C(x) \log_2 \frac{P_C(x)}{P_S(x)}.$$

Si $D_{KL}(P_C \| P_S) = 0$ alors le schéma stéganographique est dit **parfaitement sûr**

Sinon

Si $D_{KL}(P_C \| P_S) < \varepsilon$ alors dans ce cas-ci le schéma stéganographique est dit **ε- sûr**

Propriétés d'un schéma stéganographique

❑ La sécurité d'un schéma stéganographique

Point de vue pratique :

- En pratique il est difficile de définir précisément ce que c'est une distribution.

- Dans la réalité un attaquant, dont la puissance (les ressources matérielles, la capacité de calcul, le temps de calcul) est limitée, ne dispose que d'une approximation de ces distributions.

→ **Pour évaluer la sécurité d'un schéma face aux différentes formes d'attaques, auquel il doit résister, plusieurs modèles de sécurité ont été proposés (Certains modèles sont consacrés aux schémas stéganographiques à clés privées d'autres aux schémas à clés publiques ou même à un type d'adversaire particulier (passif, actif, ou malicieux)…)**

Propriétés d'un schéma stéganographique

❑ La sécurité d'un schéma stéganographique

Quelques règles de bases:

- S'assurer que le support de couverture est utilisé une seule fois, et qu'il est détruit dès son utilisation, afin d'éviter toutes les attaques par différence.
- Vérifier que la taille de la clé est suffisamment grande pou éviter les attaques exhaustives sur la clé.
- S'assurer que le processus de dissimulation du schéma stéganographique soit imperceptible à l'oeil nu pour se prémunir contre les attaques visuelles.
- De mêmes essayer de préserver au mieux la statique originale du support hôte.

Propriétés d'un schéma stéganographique

- **La sécurité d'un schéma stéganographique**
- **La capacité d'insertion**

C'est le nombre maximal de bits qui peuvent être cachés dans le médium de couverture.

Elle est souvent calculée par la capacité d'insertion relative définie par :

$$\frac{\log_2 |\mathcal{M}(x)|}{n}.$$

Où $|M|$ est le nombre de messages possibles, et $x = (x_1,\ldots, x_n)$ 2 le médium de couverture hôte composé de n éléments.

Propriétés d'un schéma stéganographique

- **La sécurité d'un schéma stéganographique**
- **La capacité d'insertion**
- **L'efficacité d'insertion :**

  le nombre de bits du message secret insérés par unité de distorsion. Autrement dit, c'est le nombre de bits de message insérés pour une modification du médium de couverture

$$e = \frac{E_x[\log_2|\mathcal{M}(x)|]}{E_{x,m}[D(x,y)]},$$

avec E l'espérance mathématique, et $D(x,y)$ la distorsion entre le support de couverture x et le stégo-support y, causée par l'insertion du message m.

Ce critère a été utilisé durent les années 2000, car on estimait que la détectabilité d'un schéma est liée au nombre de modifications (**moins il y a de modifications, moins le schéma est détectable --> Faux** .

Propriétés d'un schéma stéganographique

- **La sécurité d'un schéma stéganographique**
- **La capacité d'insertion**
- **L'efficacité d'insertion**
- **La capacité stéganographique :**

C'est le nombre maximal de bits qui peuvent être modifiés dans un médium de couverture de façon que la probabilité de détection soit insignifiante. C'est une mesure qui est strictement plus petite que la capacité d'insertion. En pratique déterminer la quantité d'information, qui peut être cachée dans le support hôte, d'une manière totalement sûre, est une tâche extrêmement difficile, et ce en raison du manque de modèles statistiques précis représentant les images numériques naturelles de nature variée et complexe.

Insertion dans le domaine spatial

Stéganographie par substitution de LSB (*LSB Replacement*)

- Le principe de cette technique consiste à substituer les bits de poids faibles (les LSB) des pixels par les bits de message à insérer.
- Le sens de parcours des pixels est usuellement choisi par un parcours pseudo-aléatoire (clé secrète **k**)

Message (M): **0 0 1 10 1 0 0**

Image en niveau de gris

01001011	00101011	10010100	11101000
10011111	01110110	00110111	00110101

Insertion dans le domaine spatial

Stéganographie par substitution de LSB (*LSB Replacement*)

- Le principe de cette technique consiste à substituer les bits de poids faibles (les LSB) des pixels par les bits de message à insérer.
- Le sens de parcours des pixels est usuellement choisi par un parcours pseudo-aléatoire (clé secrète **k**)

Message (M): **0 0 1 10 1 0 0**

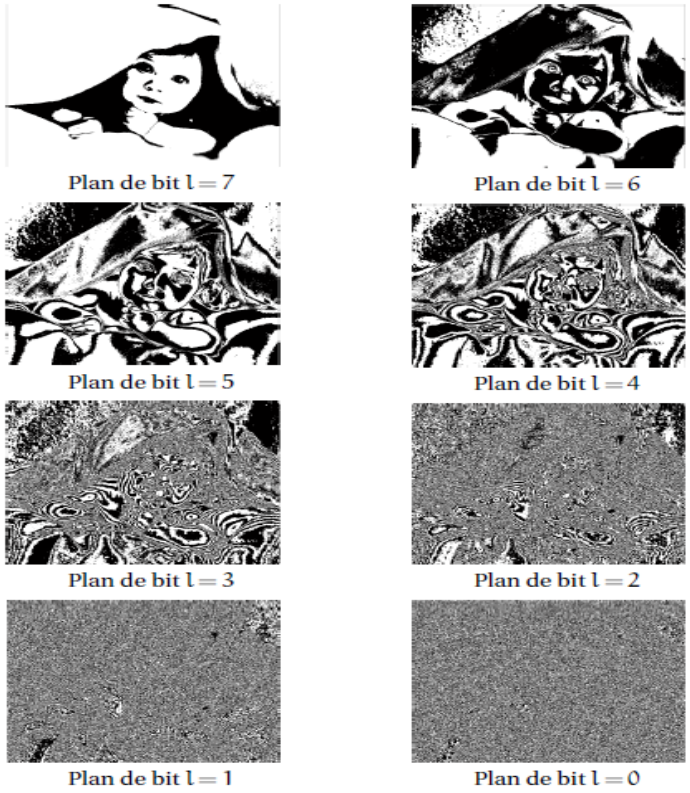
Image en niveau de gris

01001010	00101010	10010101	11101001
10011110	01110111	00110110	00110100

Insertion dans le domaine spatial

Stéganographie par substitution de LSB (*LSB Replacement*)

(a) Image en niveau de gris



(b) Images binaires des plans de bit

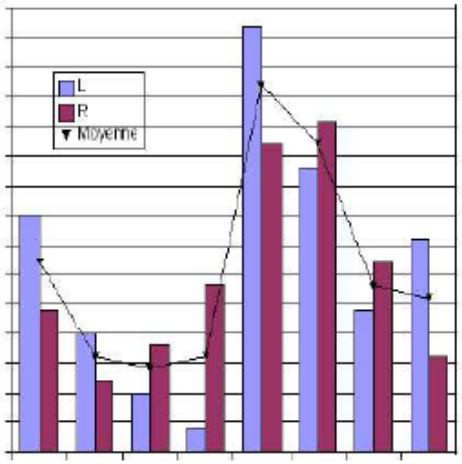
Insertion dans le domaine spatial

Stéganographie par substitution de LSB (*LSB Replacement*)

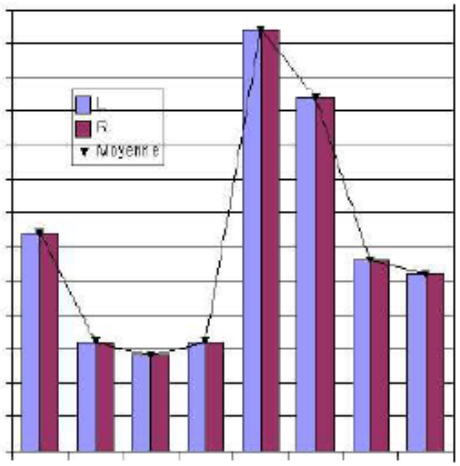
\+ Facilité d'implémentation et complexité de calcul faible.

\- Une technique stéganographique très facilement attaquable, car elle altère considérablement la distribution statistique du support hôte (attaque du χ^2).

Exemple illustratif



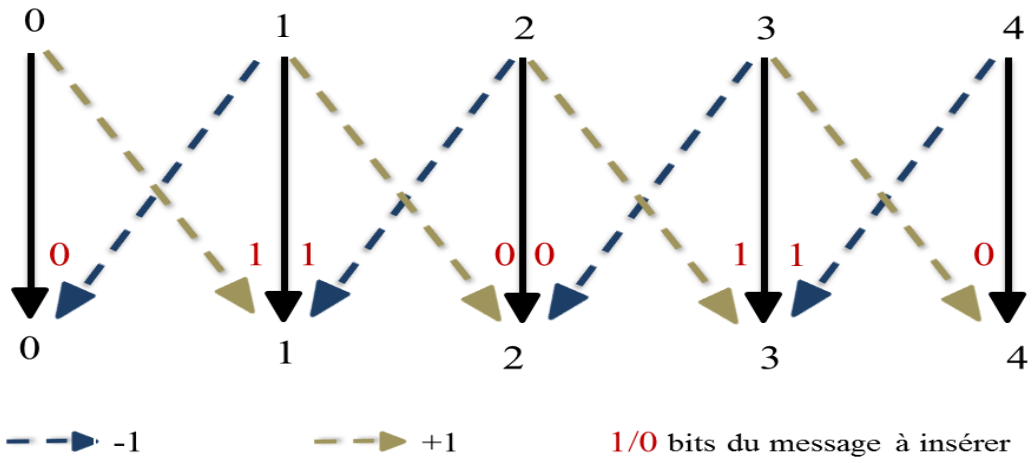
(a) Image de couverture



(b) Image stéganographiée

Insertion dans le domaine spatial

Stéganographie par correspondance de LSB (LSB Matching ou ± 1)



+ La méthode de stéganographie par correspondance des LSB n'altère pas la distribution statistique du premier ordre du support hôte. Ainsi toutes les attaques ciblés sur la statistique du premier ordre sont inefficaces.

Insertion dans le domaine transformé

- Les images échangées sur Internet sont le plus souvent compressées avec pertes au format **JPEG** ou **JFIF**.
- La plupart des méthodes de stéganographie qui opèrent dans un domaine transformé, sont **des variantes des méthodes stéganographiques spatiales.**

Insertion dans le domaine transformé

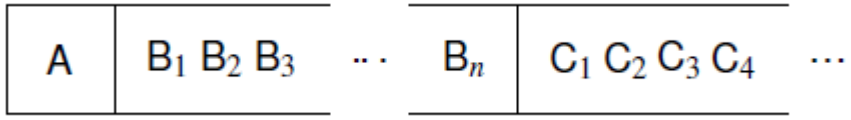
L'algorithme Jsteg (Derek Upham, 1998) :

MÉTHODE D'INSERTION DE L'ALGORITHME JSTEG :

1. Chaîne à insérer mise au bon format.
2. **Début de la compression JPEG** de l'image de couverture.
   Arrêt après l'étape de quantification.
3. **Substitution des LSB** des coefficients par les bits de la chaîne à insérer dans les coefficients DCT $\neq \{0, 1\}$.
4. Fin de la compression JPEG.

Insertion dans le domaine transformé

L'algorithme Jsteg (Derek Upham, 1998) :

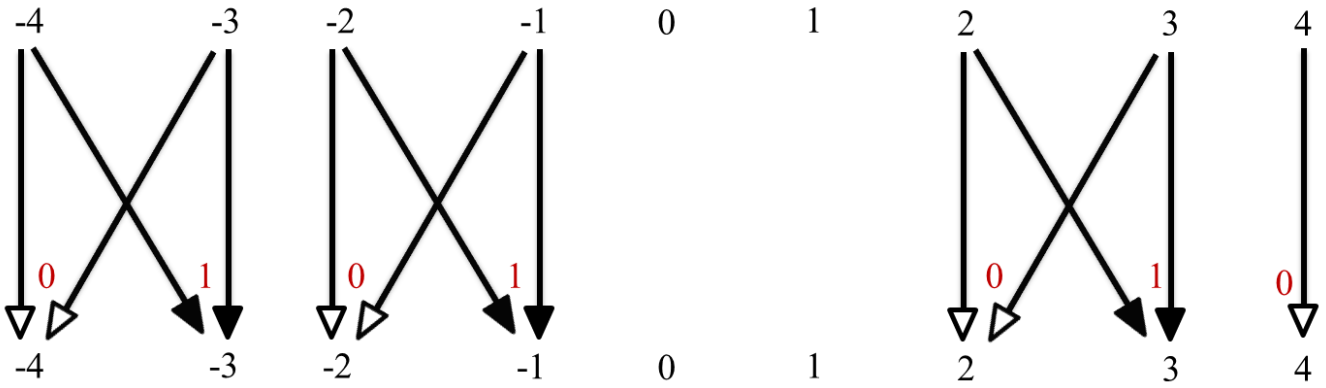


Format de la chaîne à insérer avec Jsteg.

- **A : est codé sur 5 bits, et renseigne la longueur (en bits) du champ B.**
- **B : représente une suite de n bits ϵ {0, 1} qui exprime la taille (en octets) du fichier à insérer.**
- **C : représente les bits du message à insérer.**

Insertion dans le domaine transformé

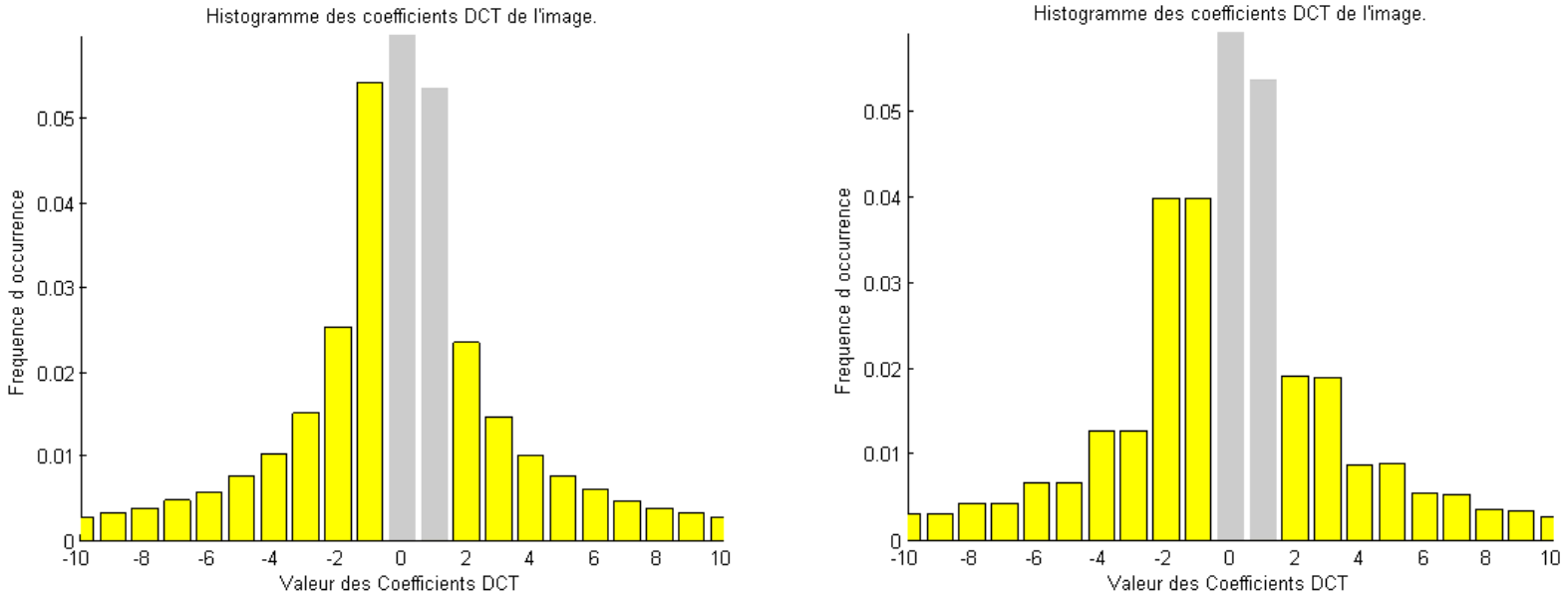
L'algorithme Jsteg (Derek Upham, 1998) :



Modification des coefficients DCT pour l'algorithme Jsteg
(Technique d'insertion par substitution, avec des valeurs omises 0 et 1)

Insertion dans le domaine transformé

L'algorithme Jsteg (Derek Upham, 1998) :



Histogramme des coefficients DCT d'une image saine (à gauche) et stéganographiée par Jsteg (à droite)

Insertion dans le domaine transformé

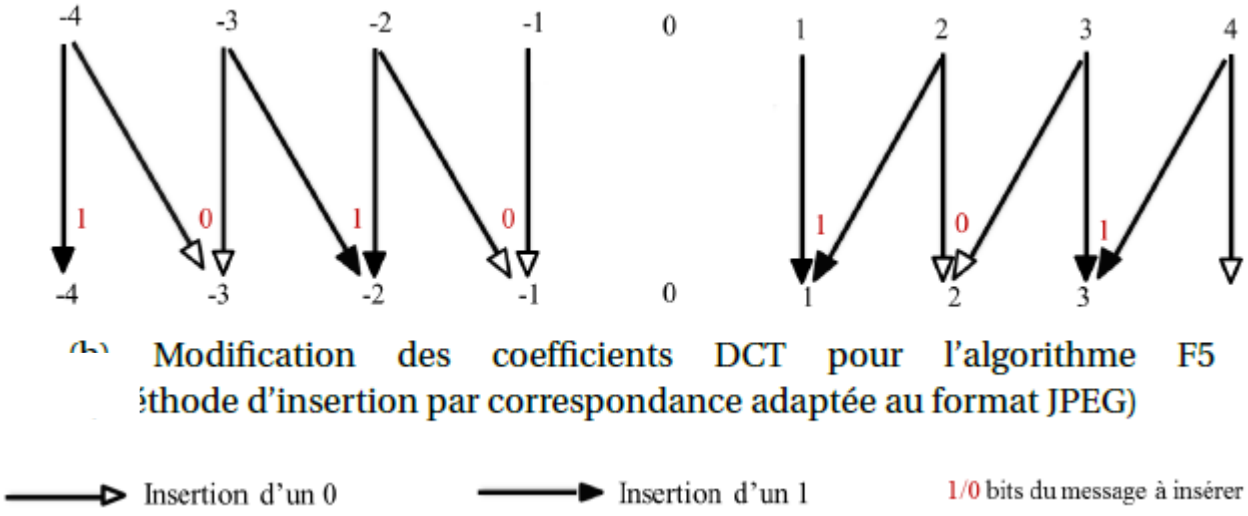
L'algorithme F5 (Westfeld, 2001) :

ALGORITHME F5 :

1. **Début de la compression JPEG** de l'image de couverture.
   Arrêt après l'étape de quantification.
2. **Initialisation du générateur de nombre pseudo-aléatoire** avec une clé.
3. **Calcul de la permutation** qui désignera le chemin à suivre pour l'insertion du message, en fonction d'un nombre aléatoire calculé précédemment et du nombre de coefficients DCT, à l'aide du générateur.
4. **Calcul des paramètres du codage de Hamming** en fonction de la capacité du medium de couverture et du message à insérer, afin de répartir au mieux les modifications.
5. **Insertion du message** avec le codage de Hamming.
6. Fin de la compression JPEG.

Insertion dans le domaine transformé

L'algorithme F5 (Westfeld, 2001) :



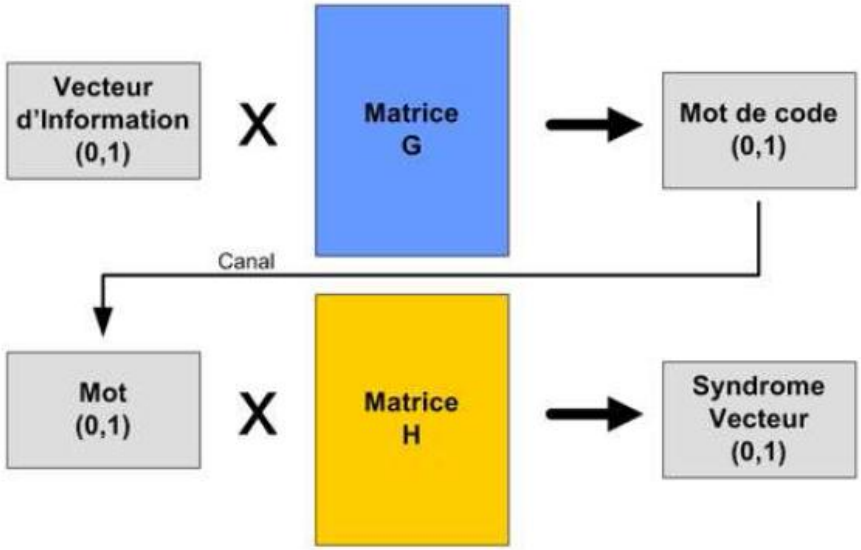
(b) Modification des coefficients DCT pour l'algorithme F5 (méthode d'insertion par correspondance adaptée au format JPEG)

Insertion d'un 0 — Insertion d'un 1 — 1/0 bits du message à insérer

Insertion dans le domaine transformé

L'algorithme F5 (Westfeld, 2001) :

- Pour augmenter l'efficacité d'insertion, L'algorithme F5 introduit pour la première fois le concept de la **technique de Matrix Embedding** pour l'insertion (Utilisation des codes de Hamming).



Insertion dans le domaine transformé

L'algorithme F5 (Westfeld, 2001) :

De manière plus formelle, le but recherché de la technique de *matrix embedding* est de communiquer un message $m \in \mathbb{F}_q^{n-k}$ (dans le cas binaire $q = 2$) au travers d'un support $x \in \mathbb{F}_q^n$, ceci en le modifiant le moins possible. Le principe est de modifier le support de couverture x en y, de telle sorte que :

$$Hy = m,$$

avec $H \in \mathcal{M}_{n-k,n}$ la matrice de parité du code. La transformation du vecteur $x \in \mathbb{F}_q^n$ en $y \in \mathbb{F}_q^n$ s'effectue alors en recherchant le vecteur de modification $e \in \mathbb{F}_q^n$:

$$y = x + e.$$

$$H(x+e) = m \quad \Leftrightarrow \quad He = m - Hx.$$

L'objectif étant donc de trouver le vecteur de modification e, celui avec le minimum de modifications possibles

Insertion dans le domaine transformé

Exemples de codes de Hamming:

Les codes de Hamming sont des codes paramétrés par un entier $p \in \mathbb{N}^+$, tel que, $n = 2^p - 1$ et $k = 2^p - 1 - p$.

Le syndrome $m - Hx = He$ représentant le message secret est de longueur $n - k = p$.

La matrice de parité du code est une matrice de $H \in \mathcal{M}_{p,2^p-1}$ formée de colonnes en binaire représentant les premiers entiers de $(1 \text{ à } 2^p - 1)$.

Insertion dans le domaine transformé

Exemples de codes de Hamming:

Pour $p = 3$, et $m = (1,0,1)$ le message que nous souhaitons insérer dans le support hôte $x = (0,1,1,1,0,0,1)$, la matrice de parité H est donc sous la forme suivante :

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

L'objectif est de trouver le vecteur $e = (e_1, e_2, e_3, e_4, e_5, e_6, e_7)$ **tel que** $H(x+e) = m$.

Insertion dans le domaine transformé

Exemples de codes de Hamming:

$$m - Hx = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

7ème colonne de H

Le vecteur de modification est alors sous la forme : $e = (0,0,0,0,0,0,1)$.

le support $x = (0,1,1,1,0,0,1)$ est alors transformé en $y = x + e = (0,1,1,1,0,0,0)$.

Insertion dans le domaine transformé

L'algorithme nsF5 (Fridrish et al., 2005) :

- L'algorithme nsF5 une solution efficace pour permettre aux deux parties (l'émetteur et le récepteur) de communiquer des informations secrètes. Il utilise **les codes à codes à papier mouillé** pour l'insertion du message secret.
- L'astuce consiste à figer certain sites sensibles à l'insertion pour éviter le risque de détectabilité.

Principe des méthodes adaptatives

Minimisation d'impact d'insertion :

Problème

Trouver le moyen de transmettre m bits dans un médium de couverture à n éléments, tout en effectuant le moins possible de distorsion.

Solution

- Modéliser l'impact d'insertion par une fonction de distorsion, souvent additive : $D(\mathbf{X}, \mathbf{Y}) = \| \mathbf{X} - \mathbf{Y} \|_\rho = \sum_{i=1}^{n} \rho_i \mid x_i - y_i \mid$.
- Trouver le stégo objet qui minimise la fonction de distorsion D sous la contrainte d'un payload fixe : $\mathbf{Y} = Emb(\mathbf{X}, m) = \arg\min D(\mathbf{X}, \mathbf{Y})$.

J.J. Fridrich, and T. Filler
Practical Methods for Minimizing Embedding Impact in Steganography. In SPIE. San Jose, CA. January 29-February 1 2007.

Principe des méthodes adaptatives

L'algorithme HUGO (Pevný et al., IH 2010) :

Calcul d'une carte de distorsion minimal, en se basant sur une mesure de détectabilité ($\rho_i \in [0,\infty[$) définie comme étant la différence entre les vecteurs caractéristiques SPAM tirés de l'image de couverture et de l'image stégo.

$$D(x,y) = \|x-y\|_\rho = \sum_{i=1}^{n} \rho_i \ |x_i - y_i| = \sum_{i=1}^{n} \rho(x, y_i x)\ |x_i - y_i| \qquad / n : NB\ pixels$$

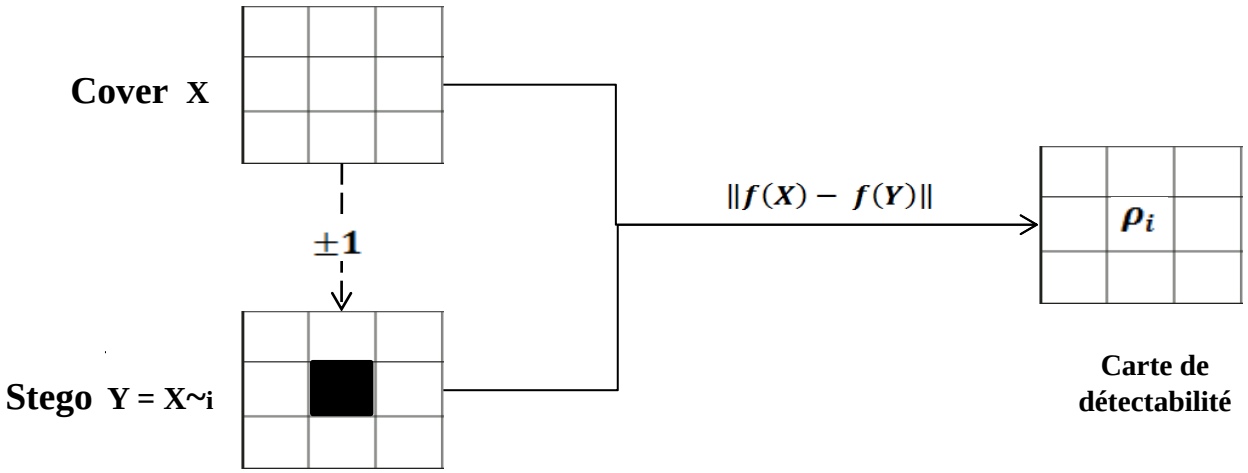
$$\rho(x,y) = \|f(x) - f(y)\| = \sum_{j=1}^{d} w_j \left|f_j(x) - f_j(y)\right| \qquad / d : NB\ features$$

Principe des méthodes adaptatives

L'algorithme HUGO (Pevný et al., IH 2010) :

$$D(x,y) = \|x-y\|_\rho = \sum_{i=1}^{n} \rho_i \ |x_i - y_i| = \sum_{i=1}^{n} \rho(x, y_i x) \ |x_i - y_i| \qquad / \ n : NB\ pixels$$

$$\rho(x,y) = \|f(x) - f(y)\| = \sum_{j=1}^{d} w_j \ |f_j(x) - f_j(y)| \qquad / \ d : NB\ features$$



Principe des méthodes adaptatives

L'algorithme HUGO (Pevný et al., IH 2010) :

Insérer le message *m* dans le média *x* tout en essayant de minimiser la fonction *D(x, y)*

1. Calcul de la carte de détectabilité:

$$\rho_i = min(\rho_{i+1}, \rho_{i-1}) \qquad /i \in \{1, \dots n\}.$$

2. Détermination des sites (pixels) les plus adaptés pour l'insertion :

ou → **Simulation de l'insertion parfaite**

ou → **Utilisation de l'approche Treillis**

3. Modification réelle des sites par **+1** ou **-1**:

Sélection de la modification qui entraine la plus petite valeur de distorsion *D*

Principe des méthodes adaptatives

L'algorithme HUGO (Pevný et al., IH 2010) :

Insérer le message *m* dans le média *x* tout en essayant de minimiser la fonction *D(x, y)*

1. Calcul de la carte de détectabilité:

$$\rho_i = min(\rho_{i+1}, \rho_{i-1}) \qquad /i \in \{1, \dots n\}.$$

2. Détermination des sites (pixels) les plus adaptés pour l'insertion :

ou → **Simulation de l'insertion parfaite**

ou → **Utilisation de l'approche Treillis**

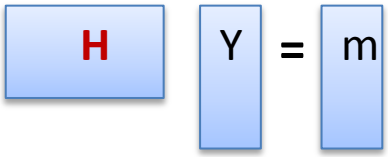
3. Modification réelle des sites par **+1** ou **-1**:

Sélection de la modification qui entraine la plus petite valeur de distorsion *D*

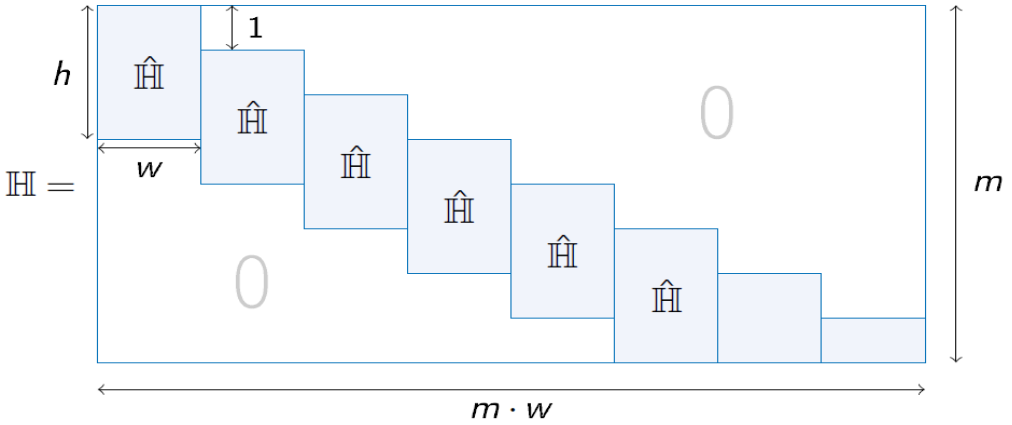
Principe des méthodes adaptatives

L'approche Treillis (Filler et al., IH 2010) :

H $\in\{0,1\}^{mxn}$: matrice de parité partagée entre l'émetteur et le récepteur

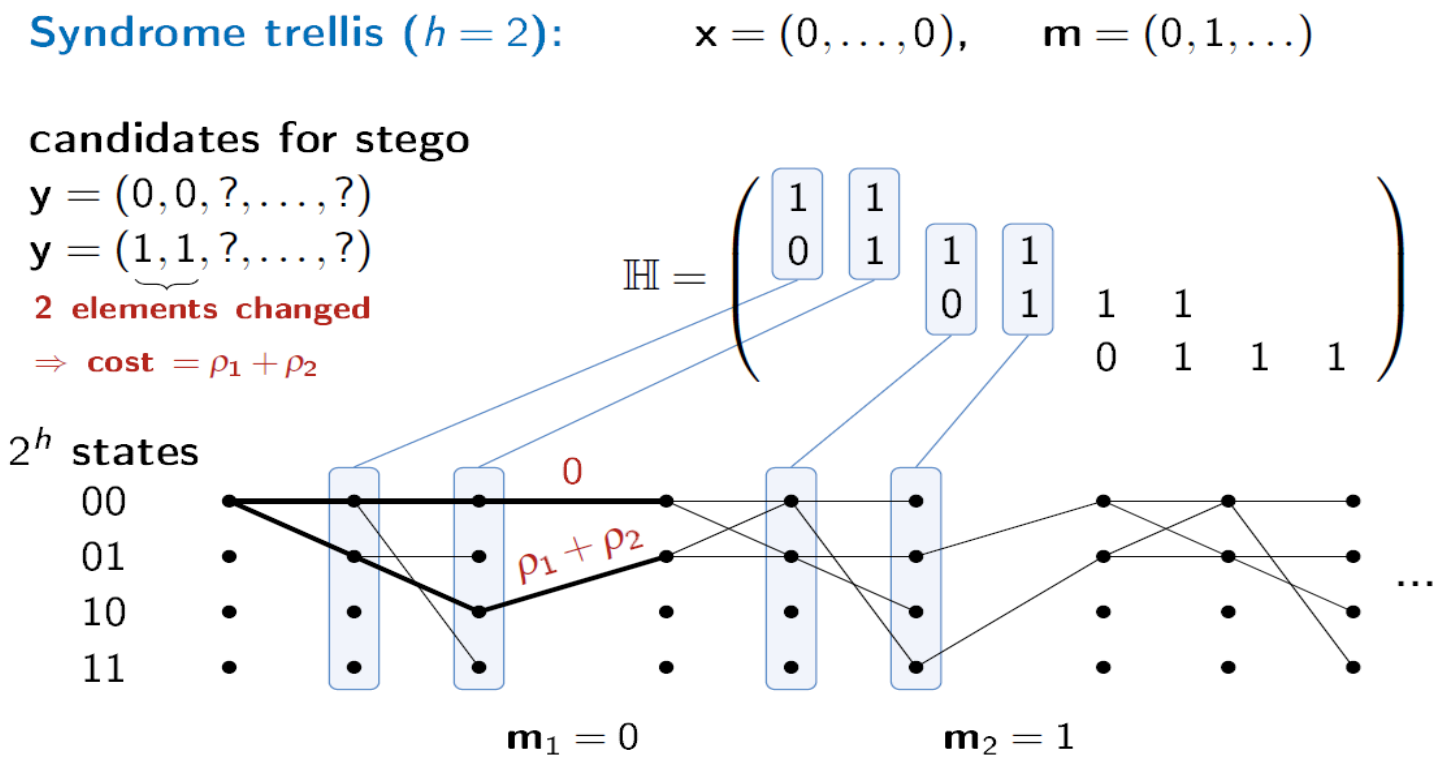


Le calcul de la matrice de parité $\mathbb{H}$ se fait à travers l'utilisation d'une sous matrice $\hat{\mathbb{H}}$ de taille (h x w):
$h \in \{1,\ldots,15\}$ et $w = 1/\alpha$.



Principe des méthodes adaptatives

L'approche Treillis (Filler et al., IH 2010) :



UNIVERSITÉ MONTPELLIER 2
SCIENCES ET TECHNIQUES

La stéganographie :

Introduction générale
Propriétés d'un schéma stéganographique

Méthodes d'insertion usuelles:

Insertion dans le domaine spatial
Insertion dans le domaine transformé

Méthodes d'insertion adaptatives:

Le principe des algorithmes adaptatifs
L'algorithme HUGO
L'approche treillis (STC)

La stéganalyse :

Définition de la stéganalyse
Principaux scénarios en stéganalyse
La stéganalyse ciblée
La stéganalyse aveugle
La stéganalyse sous d'autres angles

CONTACT :
Email : kouider@lirmm.fr
http://www2.lirmm.fr/~kouider

Stéganographie et stéganalyse

La stéganalyse

La stéganalyse (Steganalysis) - Définition

La stéganalyse

La stéganalyse a pour objectif de détecter la présence de données dissimulées à l'aide d'un algorithme stéganographique. Elle est la discipline duale de la stéganographie.

Image Hôte

≈

Image Stego

La stéganalyse (Steganalysis) – Types de gardiens

- **Gardien passif** : gardien qui se contentent d'observer le trafic entre Alice et Bob .

- **Gardien actif :** Ce gardien va essayer d'apporter quelques modifications sur le médium (Compression, filtrage…) dont le but de détruire le processus stéganographique s'il existe .

- **Gardien malicieux :** Gardien qui va essayer de comprendre la technique stéganographique et extraire le message, dont le but de le contourner pour ses propres fins.

La stéganalyse (Steganalysis) – Types de gardiens

- **Gardien passif** : gardien qui se contentent d'observer le trafic entre Alice et Bob .

- **Gardien actif :** Ce gardien va essayer d'apporter quelques modifications sur le médium (Compression, filtrage…) dont le but de détruire le processus stéganographique s'il existe .

- **Gardien malicieux :** Gardien qui va essayer de comprendre la technique stéganographique et extraire le message, dont le but de le contourner pour ses propres fins.

La stéganalyse (Steganalysis) – Principaux scénarios

❑ La stéganalyse à clairvoyance :

Dans ce scénario, le stéganographe (Alice et Bob) considère que la gardienne (Eve) dispose de tous les éléments du schéma stéganographique, à l'exception de la clé secrète utilisée lors de l'insertion (**P_C, P_S, l'algorithme, la quantité de bits insérés**).

❑ La stéganalyse à payload inconnu :

Le scénario de stéganalyse à payload inconnu est très similaire à celui de la stéganalyse à clairvoyance, à l'exception du *payload* qui est inconnu. Dans ce scénario, la communicationd u message secret est effectuée à travers une seule image.

Test entre deux hypothèses

$$\begin{cases} H_0 : \alpha \simeq 0 \\ H_1 : \alpha \geq \alpha_0 \end{cases}$$

La stéganalyse (Steganalysis) – Principaux scénarios

❑ **La stéganalyse universelle** :

Dans le cas de la stéganalyse universelle, la communication du message secret est également réalisée à travers une seule image. Pour ce scénario, le stéganographe considère que Eve la gardienne ne connaît ni l'algorithme de stéganographie utilisé, ni la quantité de bits insérés, ni la clé stéganographique**. Eve connaît seulement la distribution des images sources P_C** ; ce qui revient en pratique à choisir entre les deux hypothèses suivantes :

$$\begin{cases} H_0 : \mathbf{x} \sim P_C & \text{la distribution de l'image } \mathbf{x} \text{ est proche d'une distribution } \textit{cover} \\ H_1 : \mathbf{x} \not\sim P_C & \text{la distribution de l'image } \mathbf{x} \text{ est différente d'une distribution } \textit{cover} \end{cases}$$

La stéganalyse (Steganalysis) – Principaux scénarios

- **La stéganalyse avec cover-source mismatch** :

Dans ce scénario, la communication du message secret est également réalisée à travers une En stéganalyse avec *cover-source mismatch*, Eve, la gardienne ne connaît que partiellement, ou pas du tout, l'origine et la distribution des images de couverture sources.

- **La stéganalyse par mise en commun (pooled steganalysis)**:

Dans ce scénario, on s'approche de l'idée d'un stéganalyseur de trafic automatique qui analyse ce qui passe sur le réseau de communication.
Deux cas de figures sont possibles pour ce scénario :

- Le premier cas consiste à voir un seul acteur malintentionné qui peut parfois envoyer des données numériques contenant des informations secrètes cachées.

- Le deuxième cas consiste à avoir plusieurs acteurs qui communiquent sur le réseau et échangent des données numériques différentes.

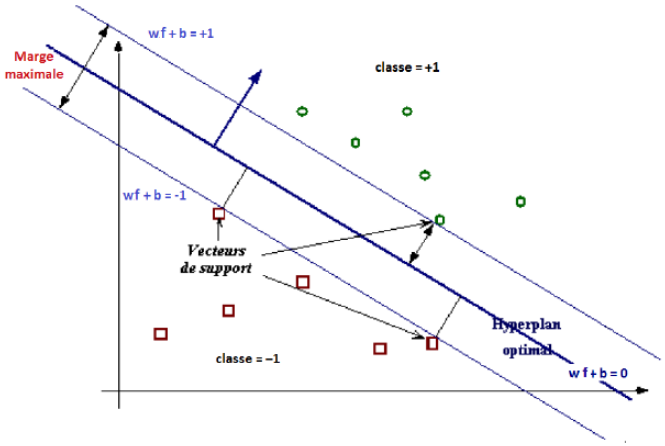
La stéganalyse (Steganalysis) – Attaques ciblées

Le principe des attaques ciblées consiste à déterminer les faiblesses de sécurité d'un algorithme particulier, en étudiant son "*implémentation*" et/ou ses "*failles statistiques*", pour pouvoir identifier la présence d'un message caché, par cet algorithme, dans un médium donné

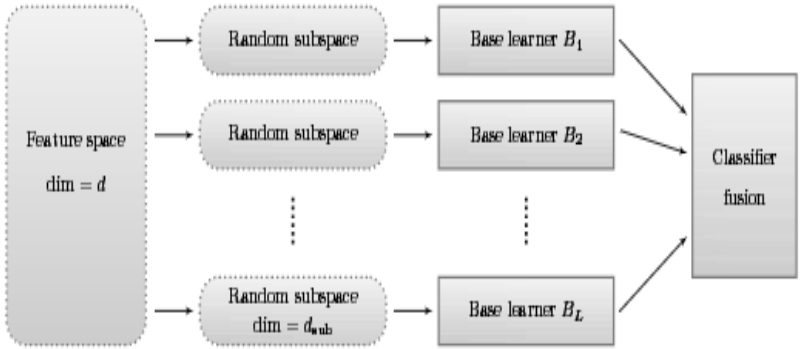
- Analyser le processus de dissimulation utilisé, pour essayer de trouver une faille particulière dans l'implémentation de l'algorithme stéganographique.
- Identifier les caractéristiques qui ont un changement prévisibles, lors de la modification du support de couverture hôte.
- Trouver le moyen d'estimer certaines caractéristiques statistiques, permettant de modéliser l'image de couverture originale, à partir de l'image stéganographiée interceptée.
- chercher une/des caractéristiques, spécifiques au format d'images, qui ont une valeur connue dans les images de couverture.

La stéganalyse (Steganalysis) – Attaques aveugles

Le principe des attaques aveugles consiste à utiliser un mécanisme d'apprentissage particulier (SVM, réseau de neurones, ensemble de classifieurs,...etc) pour détecter la présence d'un message caché au sein d'un support donné. Ces méthodes sont plus performantes que les méthodes d'attaques ciblées.



Support Vector Machin (SVM)



Ensemble de classifieurs FLD

La stéganalyse (Steganalysis) – Autres points de vue

❑ **La stéganalyse du point de vue statistique** :

- Les méthodes dites "*de détection statistiques*" reposent sur la théorie de la décision pour détecter la présence du message secret.

- Pour Juger si oui ou non un support est conteneur d'information secrète, ces méthodes établissent d'abord un modèle paramétrique définissant la nature et la statistique des images de couverture, ensuite, se fixent un critère d'optimalité donné pour choisir le test le plus adéquat au problème de la détection (test Bayésien, test minimax, ou test de de Neyman Pearson).

La stéganalyse (Steganalysis) – Autres points de vue

- **La stéganalyse du point de vue statistique**
- **La stéganalyse du point de vue de la théorie des jeux**

Le problème de stéganalyse/stéganographie peut également être abordé sous l'angle de la théorie des jeux. De manière générale, la théorie des jeux est une approche très intéressante, lorsqu'il s'agit de modéliser la stratégie de chacun des participants d'un jeu compétitif. Elle permet de prendre en compte le comportement de deux (ou plusieurs) opposants qui doivent adapter leurs stratégies en fonction d'hypothèses sur le comportement des autres adversaires dans le jeu.