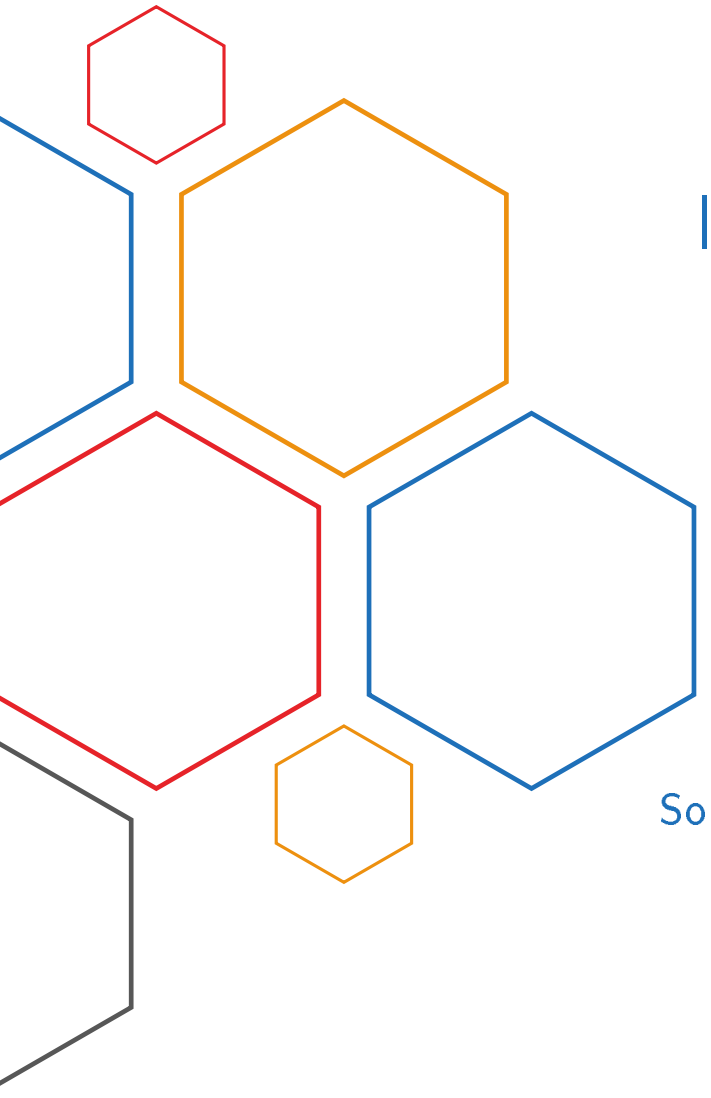


Insertion de données cachées haute capacité dans le domaine chiffré

Le 13 juillet 2018 - Journées de l'équipe ICAR

Pauline Puteaux

Sous la direction de William Puech

Sommaire

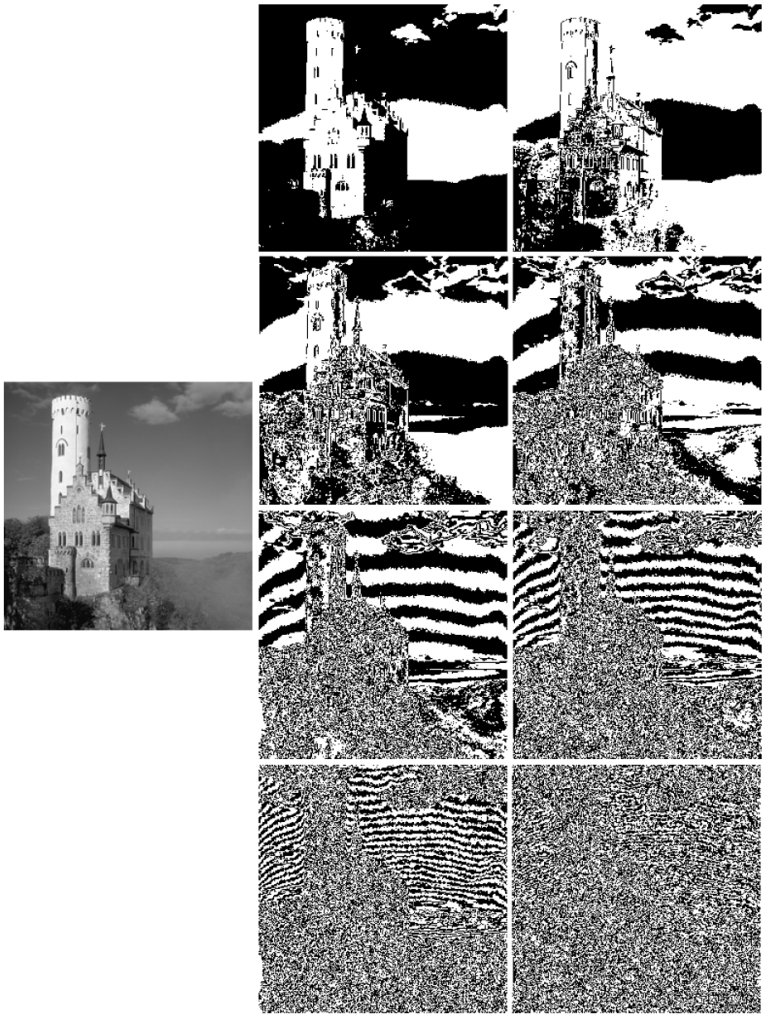
Qu'est qu'une image ?

Représentation classique

- Matrice deux dimensions
- Coefficients = pixels
- Codage sur 8 bits pour une image en niveaux de gris (valeurs de 0 à 255)

Plans binaires

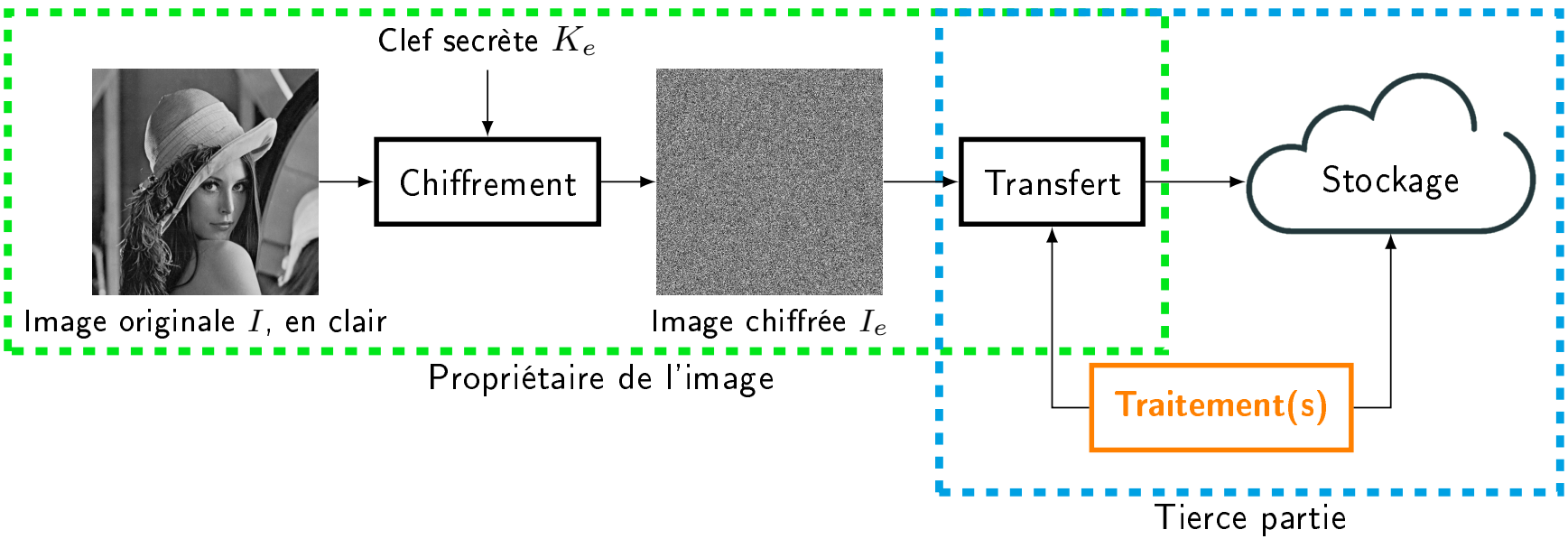
- Ensemble de bits à une position donnée dans chacun des pixels
- Bit le plus significatif : MSB
- Bit le moins significatif : LSB



Les 8 plans binaires
d'une image en niveau de gris.

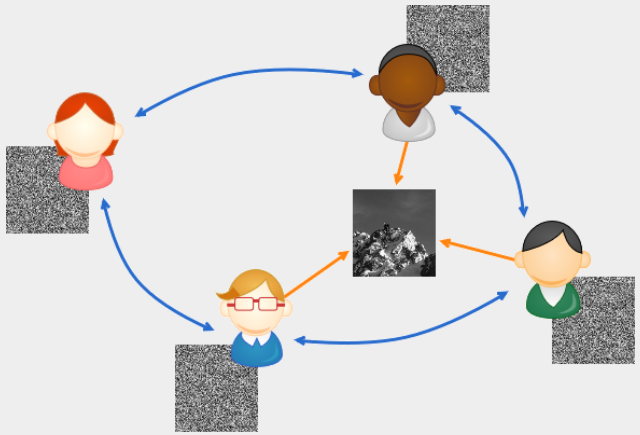
Problème de la sécurité des données multimédia

- Transfert ou archivage de ces données sous forme chiffrée
- Préservation du format
- Nécessité de les analyser et de les traiter sans la clef

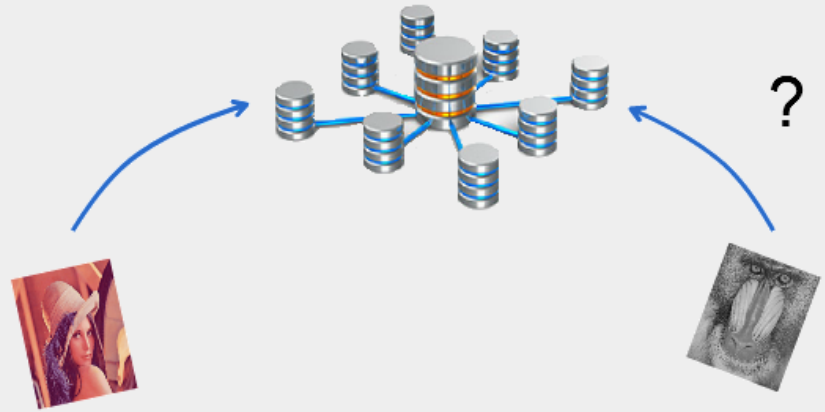


Applications

Images secrètes partagées ("Visual Secret Sharing")

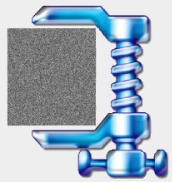


Recherche/indexation dans des BDD chiffrées

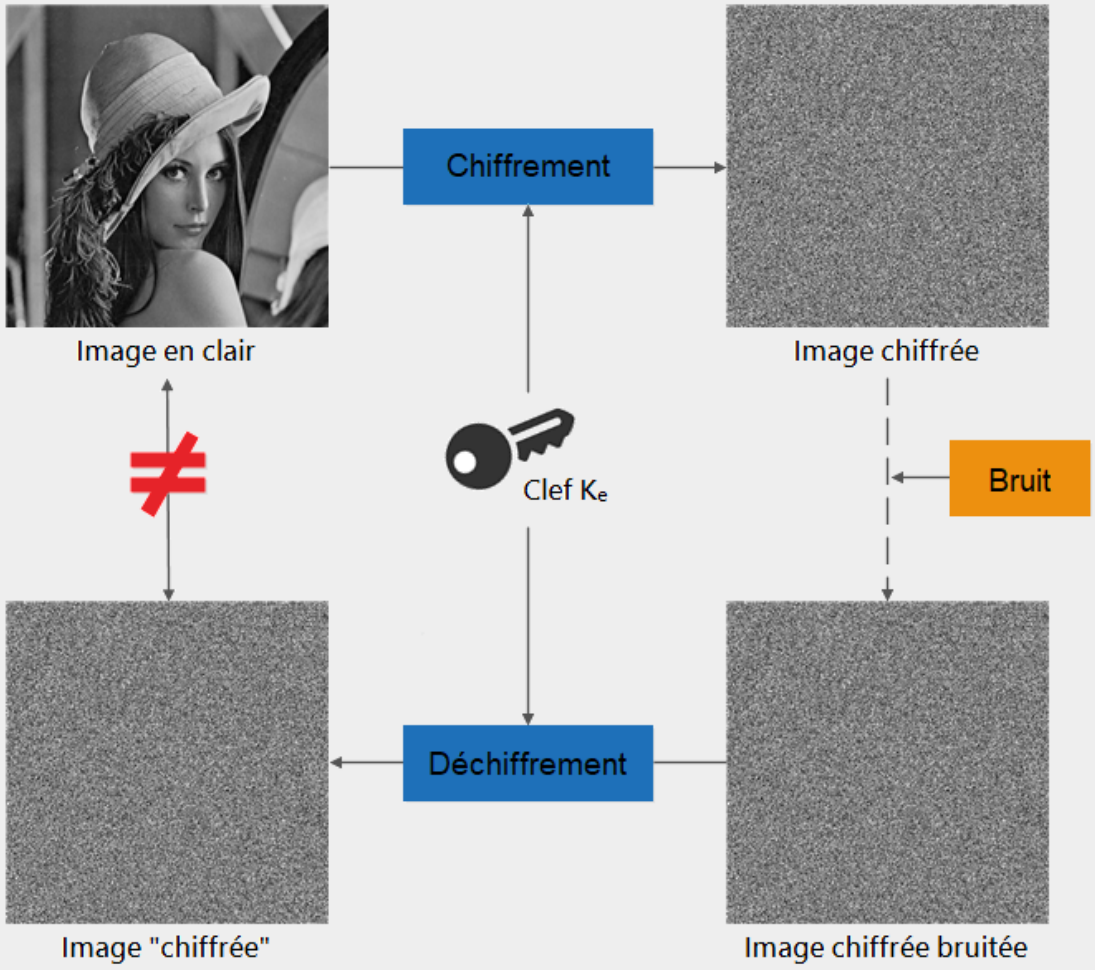


Insertion de données cachées dans le domaine chiffré

Recompression d'images crypto-compressées

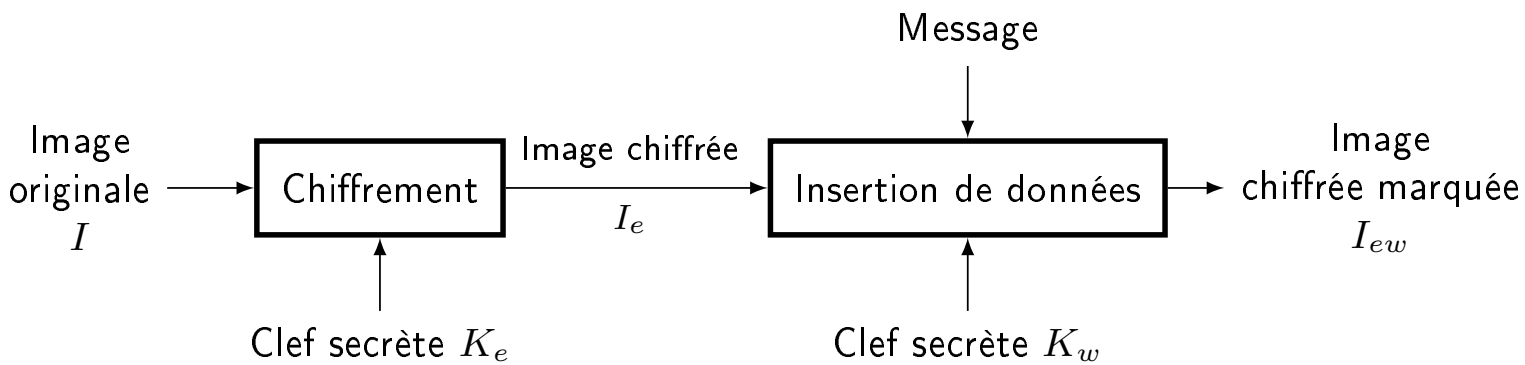


Reconstruction de l'image originale ?



Définition

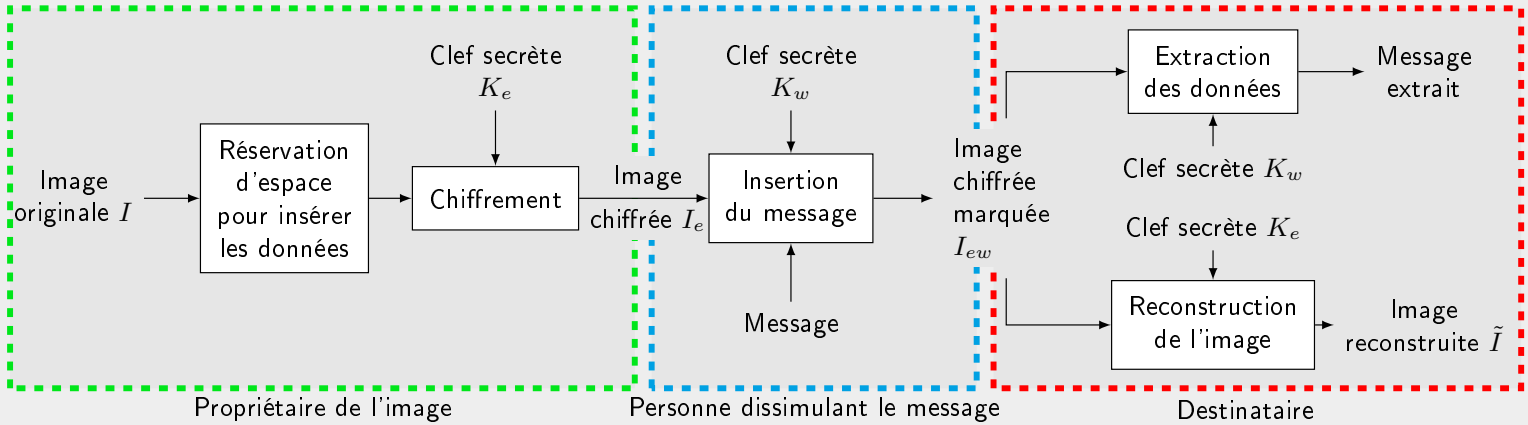
- Approche pour insérer des données secrètes dans le domaine chiffré, sans connaître le contenu de l'image originale ou la clef de chiffrement utilisée
- Phase de déchiffrement : message extrait sans erreur, image originale reconstruite sans perte
- Capacité d'insertion *vs* qualité de l'image reconstruite



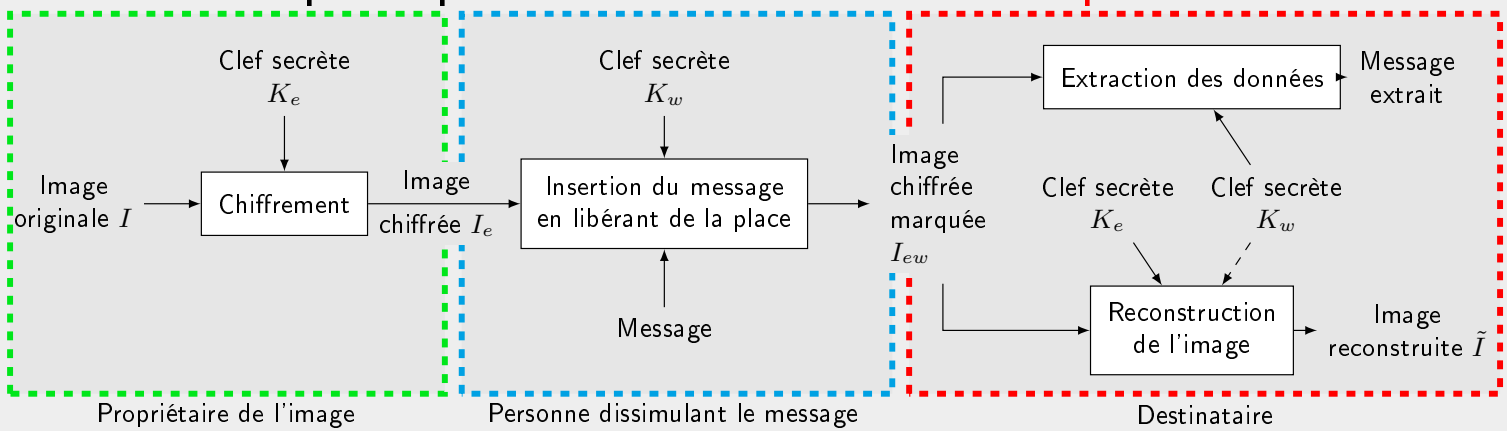
Sommaire

2 approches différentes pour l'insertion

1. Réserver l'espace pour insérer les données avant le chiffrement



2. Réserver l'espace pour insérer les données après le chiffrement



2 catégories pour l'extraction

1. Extraction des données et reconstruction de l'image conjointement
2. Reconstruction de l'image et extraction du message séparément
   - Extraction du message seulement (à l'aide de K_w)
   - Reconstruction de l'image seulement (à l'aide de K_e)
   - Reconstruction de l'image et extraction du message (avec K_w et K_e)

X. Zhang

Separable reversible data hiding in encrypted image

IEEE Transactions on Inf. Forensics and Security, vol. 7, no. 2, pp.826–832, 2012

K. Ma, W. Zhang, X. Zhao, N. Yu and F. Li

Reversible data hiding in encrypted images by reserving room before encryption

IEEE Transactions Inf. Forensics and Security, vol. 8.3, pp.553-562, 2013

W. Zhang, K. Ma and N. Yu

Reversibility improved data hiding in encrypted images

Signal Processing, vol. 94, pp.118-127, 2014

Substitution pseudo-aléatoire de bits dans l'image

- Chiffrement de l'image par bloc
- Insertion d'un ou plusieurs bit(s) dans chaque bloc chiffré
- Utilisation d'un critère statistique
- Exploitation de la différence entre un bloc en clair *vs* chiffré

Avantages

Reconstruction exacte de l'image originale (avec la clef de chiffrement seulement), extension possible à la correction d'images chiffrées bruitées

Inconvénients

Données insérées non conservées après le déchiffrement de l'image, capacité d'insertion relativement faible (si totalement réversible)

W. Puech, M. Chaumont, and O. Strauss

A reversible data hiding method for encrypted images

Electronic Imaging, International Society for Optics and Photonics, 2008, pp. 68191E

Substitution des bits les moins significatifs (LSB)

- Méthode très utilisée
- Utilisation des méthodes d'insertion de données classiques :
  - Compression
  - Décalage d'histogramme
  - Prédiction

Avantage (méthodes récentes)

Chiffrement homomorphe à l'insertion de données cachées

Inconvénients

Très faible capacité d'insertion ($\approx 0.1\ bpp$), reconstruction de l'image originale avec pertes

X. Zhang, J. Long, Z. Wang, and H. Cheng

Lossless and reversible data hiding in encrypted images with public-key cryptography

IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 9, pp. 1622–1631, 2016

Substitution des bits les plus significatifs (MSB)

- Nouvelle approche
- Exploitation de la corrélation entre un pixel et son voisinage
- Prédiction à l'aide du voisinage

Avantages

Très grande capacité d'insertion ($\approx 1\ bpp$), reconstruction exacte de l'image originale (avec la clef de chiffrement seulement)

Inconvénient

Données insérées non conservées après le déchiffrement de l'image

P. Puteaux and W. Puech

An efficient MSB prediction-based method for high-capacity reversible data hiding in encrypted images

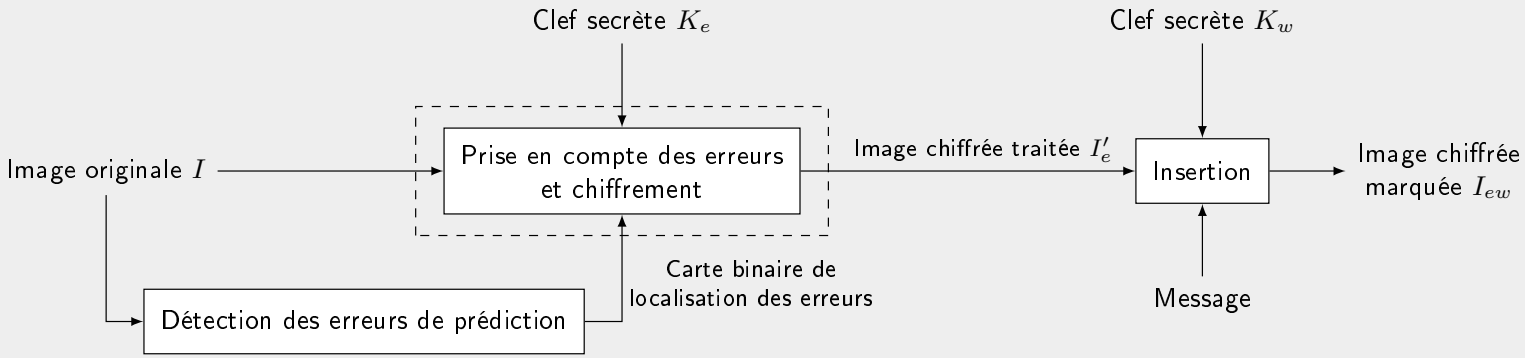
IEEE Transactions on Inf. Forensics and Security, vol. 13, no. 7, pp.1670–1681, 2018

Sommaire

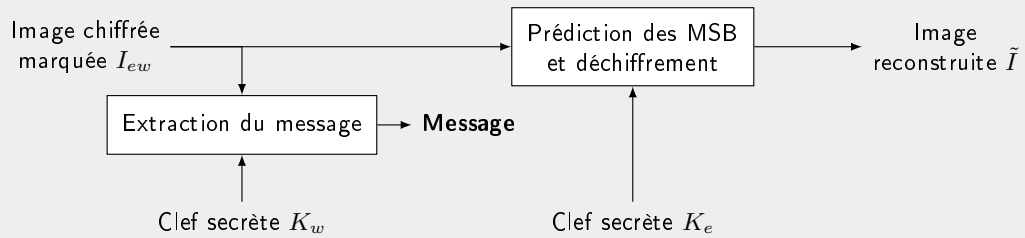
Prédiction des MSB

- Insertion effectuée en substituant les MSB des pixels
- Valeurs initiales prédites lors de la phase de reconstruction

Chiffrement de l'image et insertion du message



Reconstruction de l'image et extraction du message



Prédiction du plan MSB : 2 approches proposées

Approche$_{1bpp}$: Correction des erreurs de prédiction

- Pré-traitement de l'image originale (adaptation des pixels pour éliminer les erreurs de prédiction)
- Chiffrement de l'image pré-traitée
- Insertion de données dans tous les pixels

Résultats

- Insertion très haute capacité (capacité d'insertion = $1\ bpp$)
- Bonne qualité de l'image reconstruite (PSNR $\approx 57,4\ dB$)

P. Puteaux and W. Puech

An efficient MSB prediction-based method for high-capacity reversible data hiding in encrypted images

IEEE Transactions on Inf. Forensics and Security, vol. 13, no. 7, pp.1670–1681, 2018

Prédiction du plan MSB : 2 approches proposées

$\text{Approche}_{\text{reversible}}$: Signalement des erreurs de prédiction

- Chiffrement de l'image originale
- Signalement de l'emplacement des erreurs de prédiction (erreurs/drapeaux)
- Insertion de données dans tous les pixels où cela est possible

Résultats

- Haute capacité de dissimulation (capacité d'insertion $\approx 0,98\ bpp$)
- Image parfaitement reconstruite (PSNR $\rightarrow +\infty$)

P. Puteaux and W. Puech

An efficient MSB prediction-based method for high-capacity reversible data hiding in encrypted images

IEEE Transactions on Inf. Forensics and Security, vol. 13, no. 7, pp.1670–1681, 2018

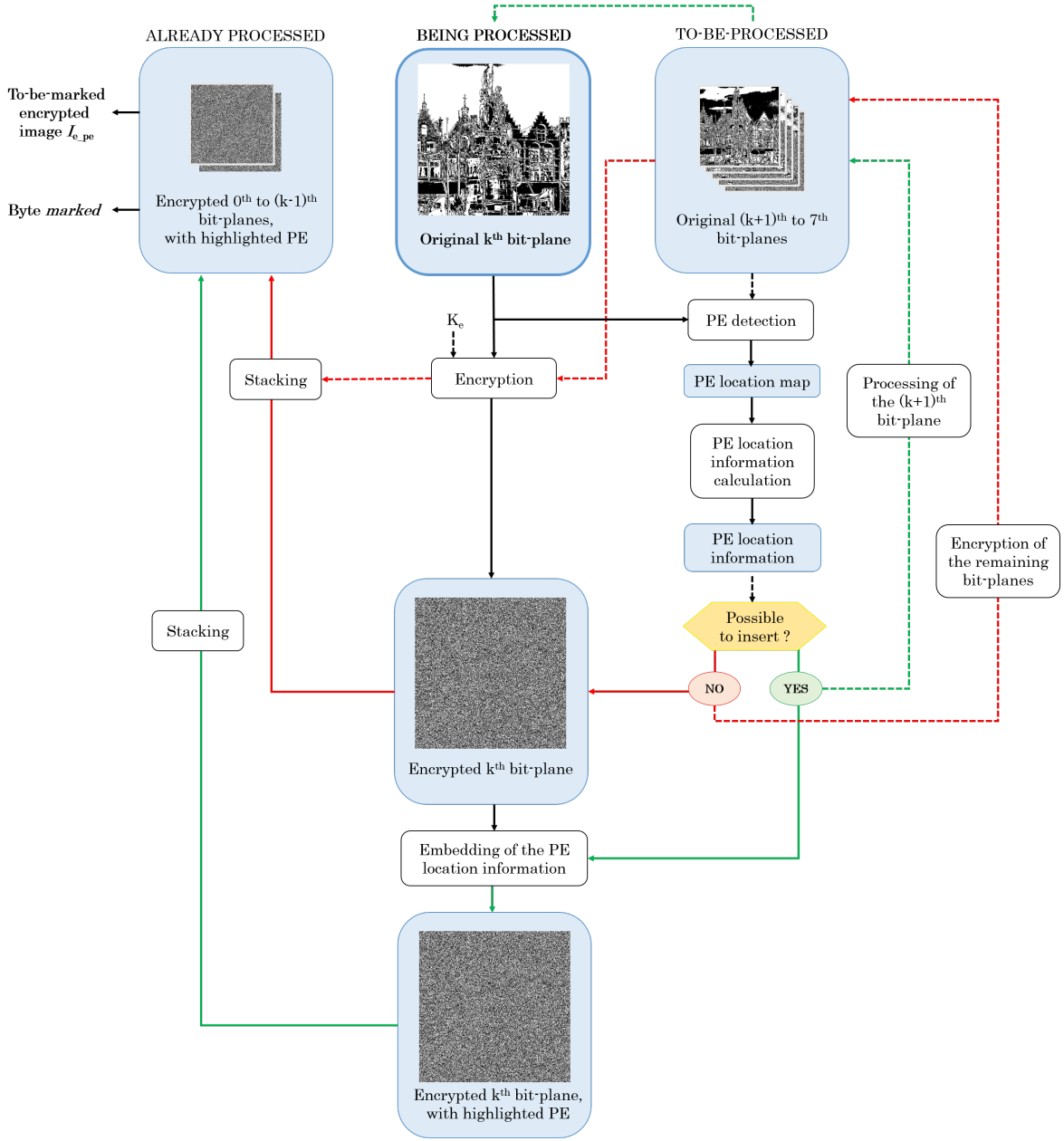
Traitement récursif de tous les plans binaires de l'image

- Approche récursive, sur tous les plans de l'image
- Le MSB n'est pas le seul bit à pouvoir être prédit
- Possibilité d'utiliser les autres plans binaires pour l'insertion
- 2 nouvelles approches définies

Extension de l'approche$_{\text{reversible}}$

- Détection des erreurs de prédiction
- Chiffrement du plan courant
- Insertion de la carte de localisation des erreurs de prédiction
- Insertion des données dans les bits restants
- Arrêt lorsque la carte de localisation des erreurs ne peut pas être insérée ou qu'un drapeau va être mal détecté
- Utilisation d'un octet pour savoir quels plans sont marqués

Approches récursives



Poupées russes : correction réversible

- Détection des erreurs de prédiction
- Stockage des valeurs et de l'emplacement des erreurs de prédiction
- Pré-traitement des plans binaires (adaptation pour signaler les erreurs de prédiction)
- Chiffrement du plan courant
- Insertion des valeurs des erreurs de prédiction
- Insertion des données dans les bits restants

Approches récursives

Image couleur originale,

Approches récursives

Image couleur originale,

et ses 3 plans Rouge, Vert, Bleu.

Composante Vert originale (plans 0 à 7),

Composante Vert originale (plans 0 à 7),

Détection des erreurs de prédiction, puis adaptation de l'image.

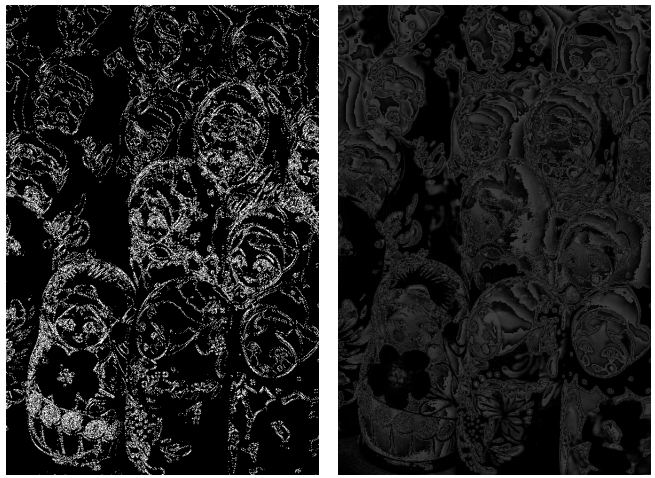
Plans 1 à 7 de l'image adaptée,

Plans 1 à 7 de l'image adaptée,

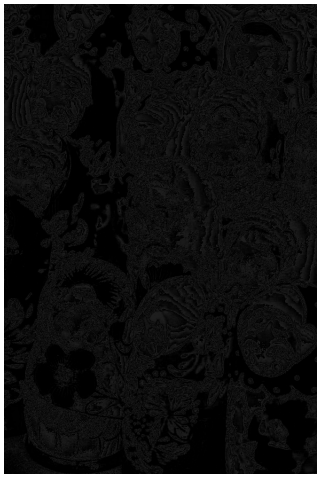

Détection des erreurs de prédiction, puis adaptation de l'image.

Plans 2 à 7 de l'image adaptée,

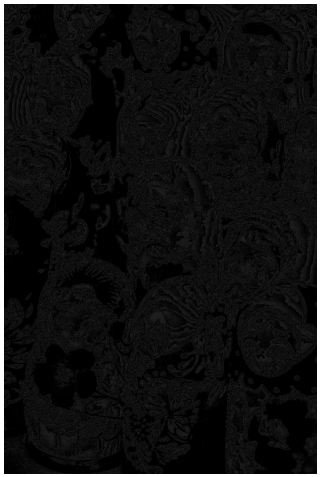
Plans 2 à 7 de l'image adaptée,



Détection des erreurs de prédiction, puis adaptation de l'image.



Plans 3 à 7 de l'image adaptée,



Plans 3 à 7 de l'image adaptée,

Détection des erreurs de prédiction $\rightarrow$ taille trop grande !
Pas d'adaptation de l'image, fin du processus récursif.

Approches récursives

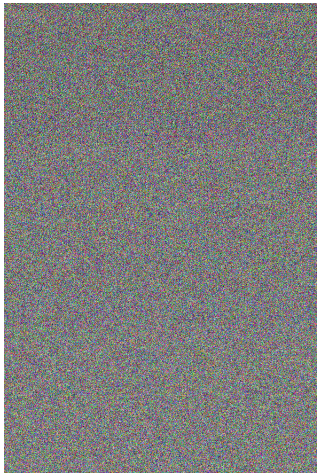


Image chiffrée marquée,

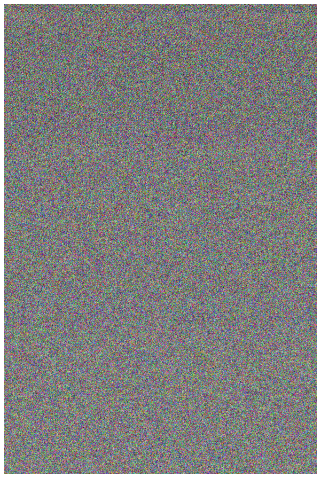
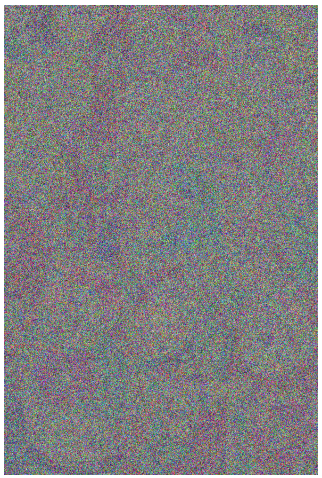
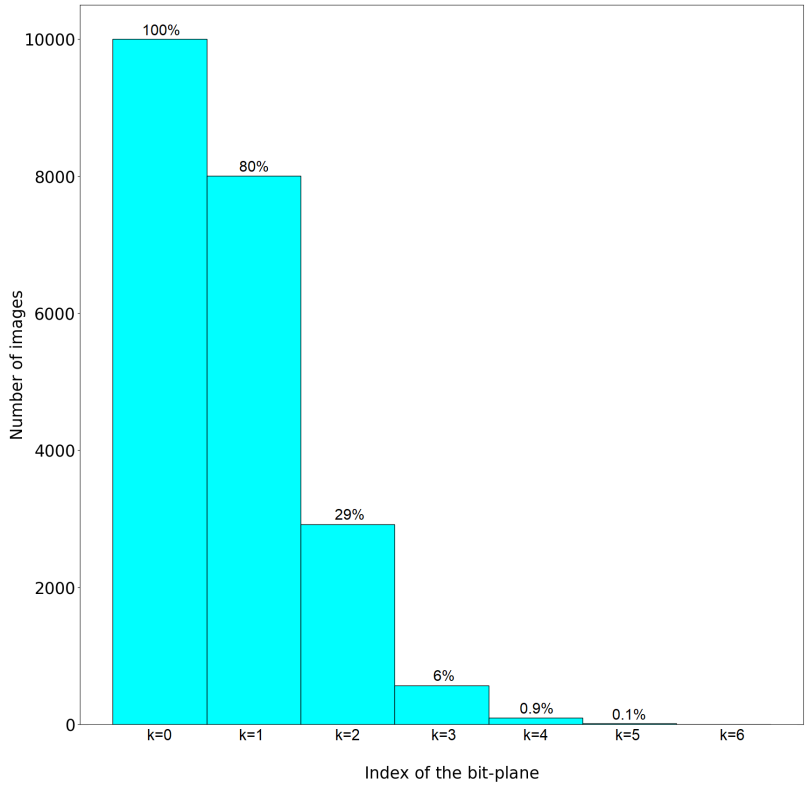


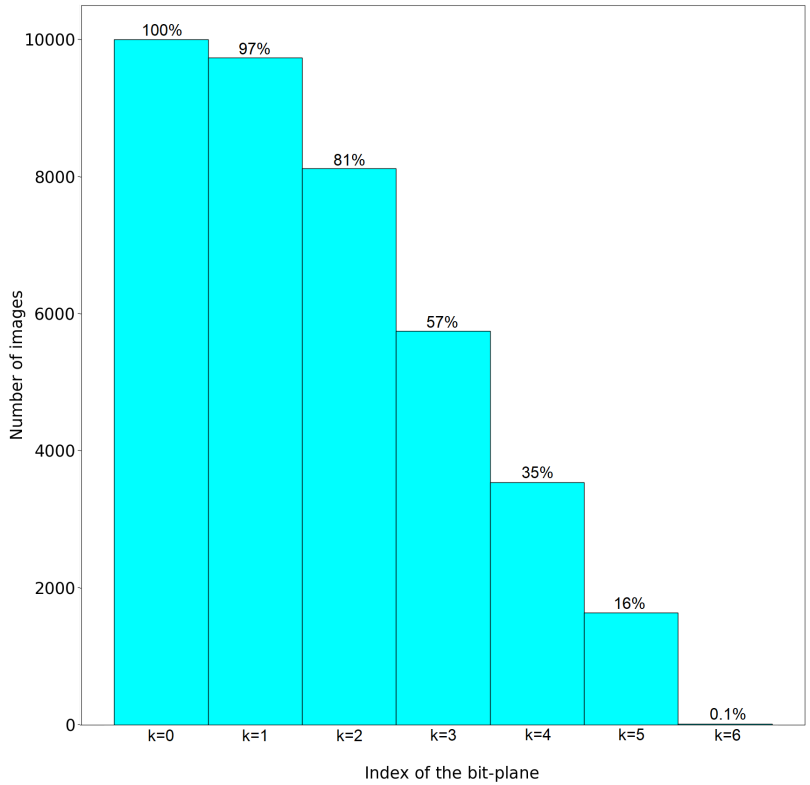
Image chiffrée marquée,



Déchiffrement naif (à gauche),
reconstruction de l'image originale (à droite).

Approches récursives



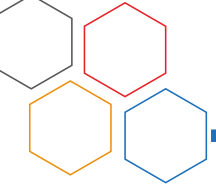


Distribution des images de la base BOWS-2 suivant la possibilité d'insérer dans le plan binaire d'indice k et pourcentages associés :

- pour l'extension de l'approche$_{\text{reversible}}$ (à gauche),
- pour la méthode des poupées russes (à droite).

Résultats

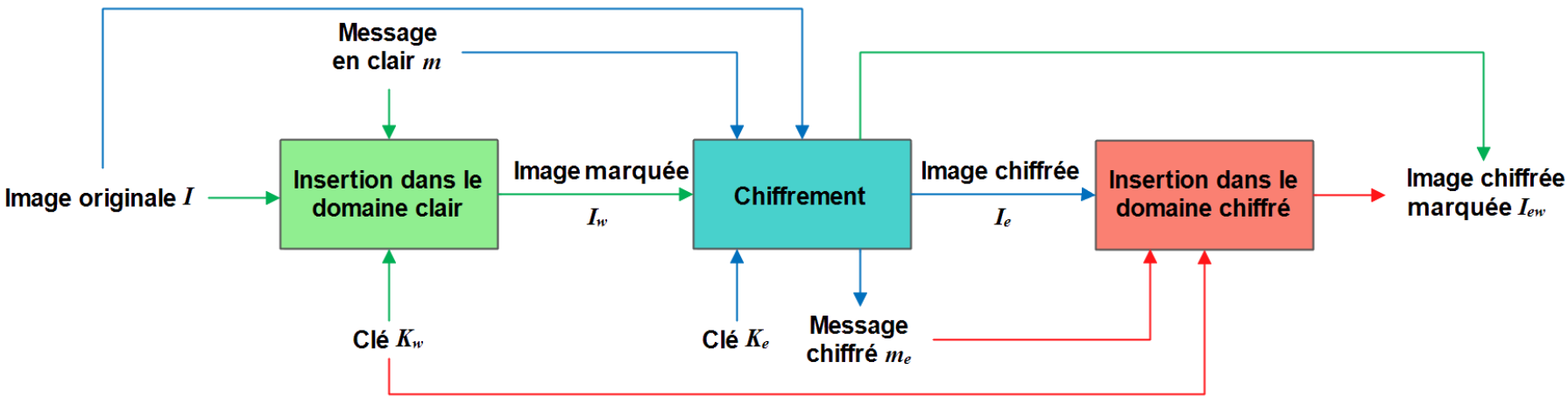
- Image parfaitement reconstruite (PSNR $\rightarrow +\infty$)
- Très haute capacité de dissimulation
  - pour l'extension de l'approche$_{\text{reversible}}$:
    - $1,75$ bpp de valeur médiane,
    - $1,84$ bpp en moyenne,
    - $5,41$ bpp dans le meilleur cas
  - pour la méthode des poupées russes :
    - $2,32$ bpp de valeur médiane,
    - $2,46$ bpp en moyenne,
    - $6,21$ bpp dans le meilleur cas



Sommaire

Propriété désirée

$$\mathcal{E}(I \oplus m) = \mathcal{E}(I) \otimes \mathcal{E}(m)$$



Problématique

- Comment évaluer le niveau de sécurité des méthodes d'insertion de données cachées dans le domaine chiffré ?
- Pas de protocole d'évaluation défini à ce jour

Définir un protocole d'évaluation

- Analyse statistique approfondie (corrélation, entropie...)
- Profils de lignes
- Résistance à la stéganalyse, évaluée en utilisant l'apprentissage supervisé
- Scénarios d'attaque, en particulier pour la cryptanalyse

Merci pour votre attention !

Questions ?