Insertion de données cachées dans le domaine chiffré

Pauline Puteaux

Le 31 mai 2018 – Pré-GdR Sécurité Informatique/GdR ISIS

Journée Traitement de données multimédia dans le domaine chiffré

Sous la direction de William Puech – LIRMM, Université de Montpellier/CNRS

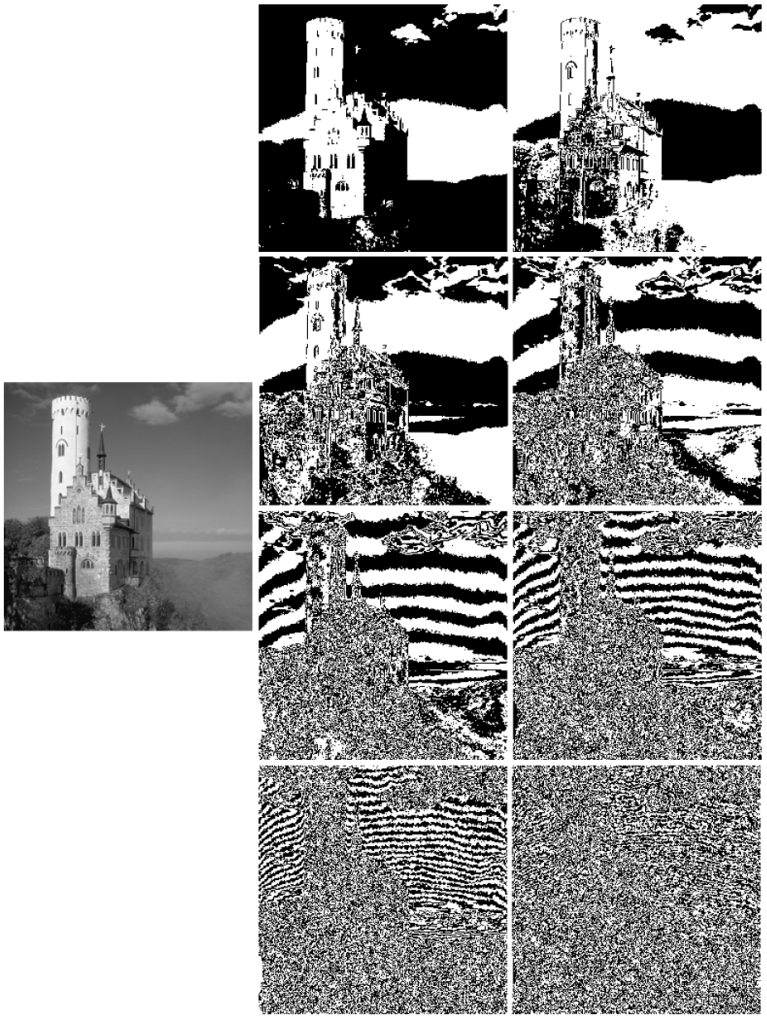
Qu'est qu'une image ?

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

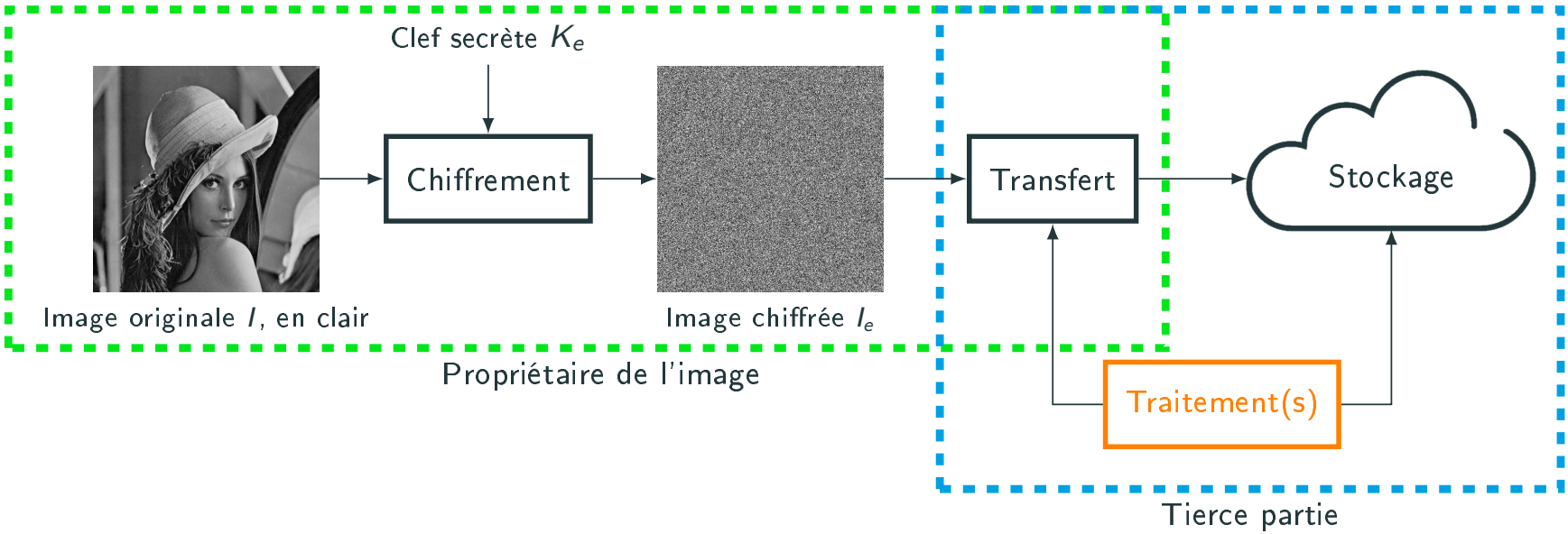


Les 8 plans binaires
d'une image en niveau de gris.

Traitement des images dans le domaine chiffré

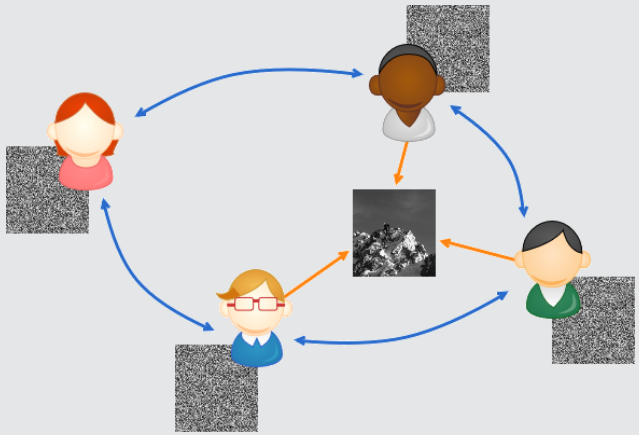
Problème de la sécurité des données multimédia

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

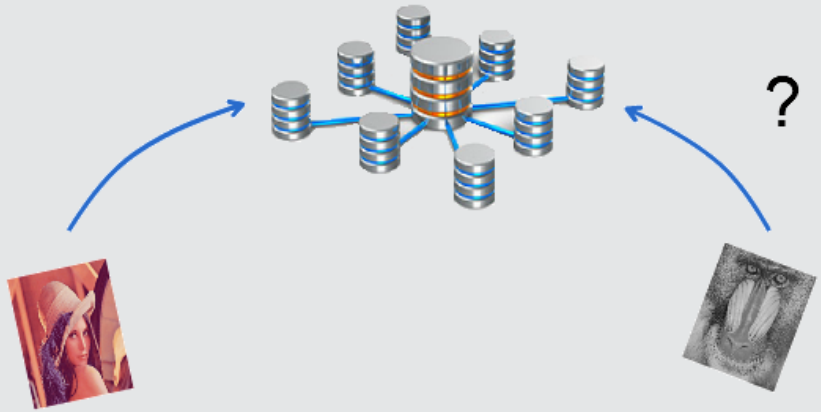


Applications

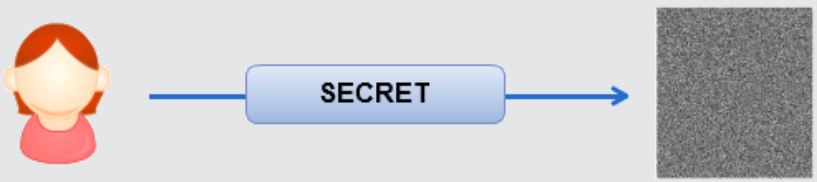
Images secrètes partagées ("Visual Secret Sharing")



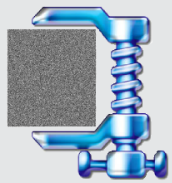
Recherche/indexation dans des BDD chiffrées



Insertion de données cachées dans le domaine chiffré

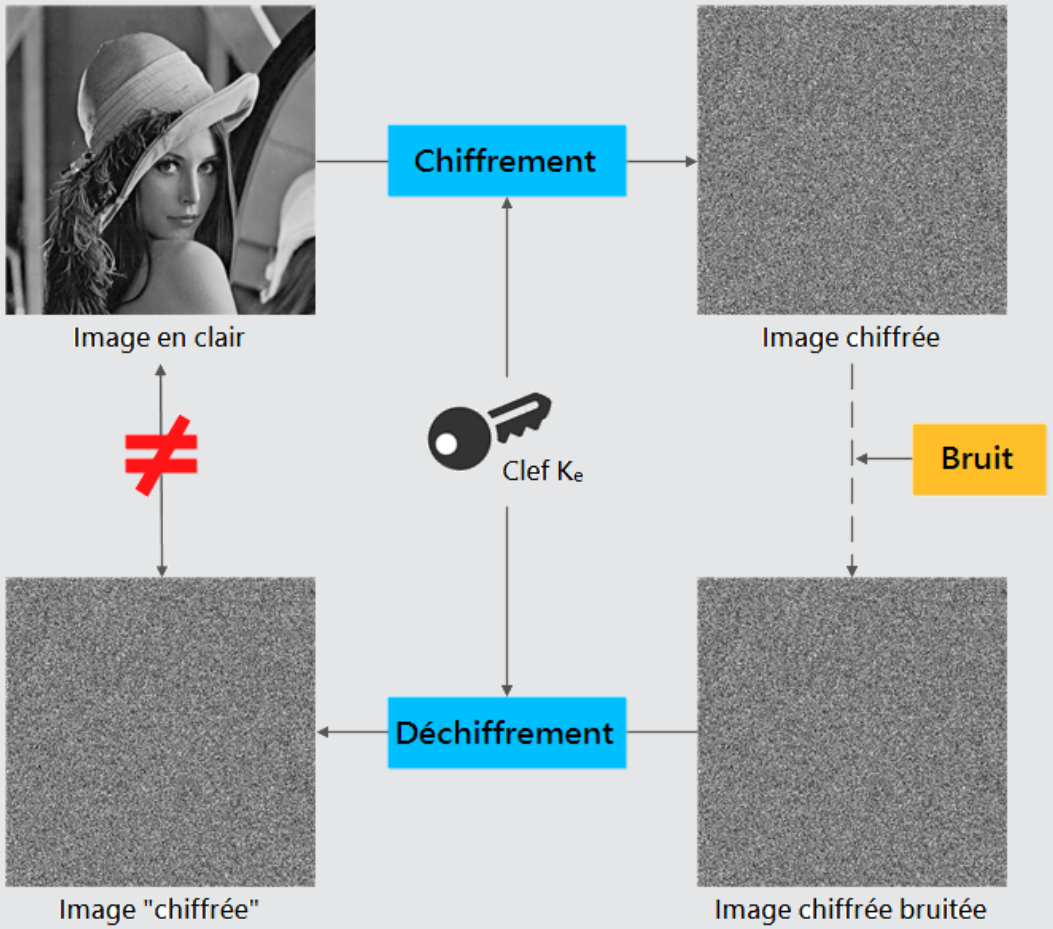


Recompression d'images crypto-compressées



Problème des traitements dans le domaine chiffré

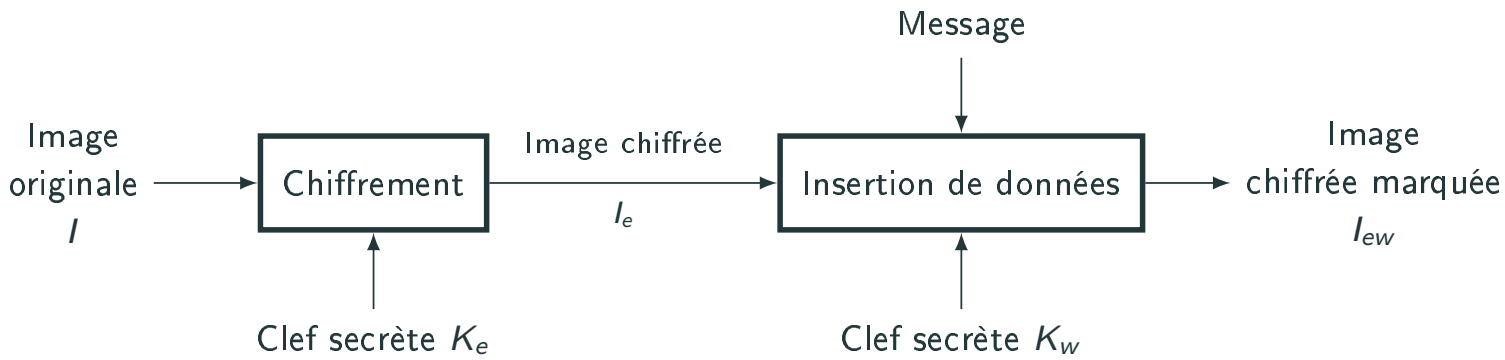
Reconstruction de l'image originale ?



Insertion de données cachées dans le domaine chiffré

Définition

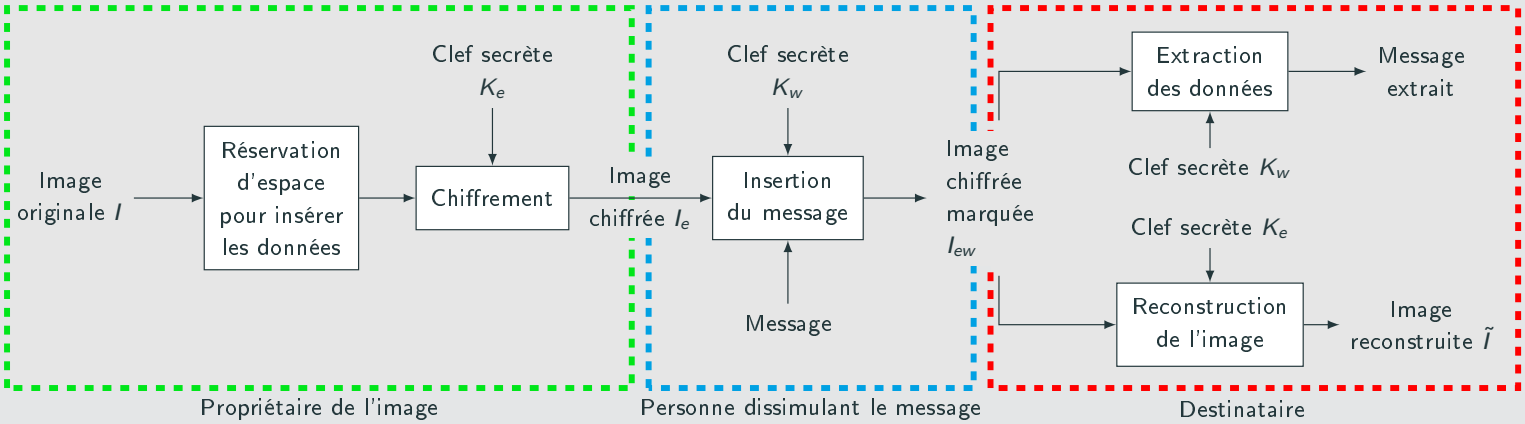
- Approche pour insérer des données secrètes dans le domaine chiffré, sans connaître le contenu de l'image originale ou la clef de chiffrement utilisée
- Phase de déchiffrement : message extrait sans erreur, image originale reconstruite sans perte
- Capacité d'insertion *vs* qualité de l'image reconstruite



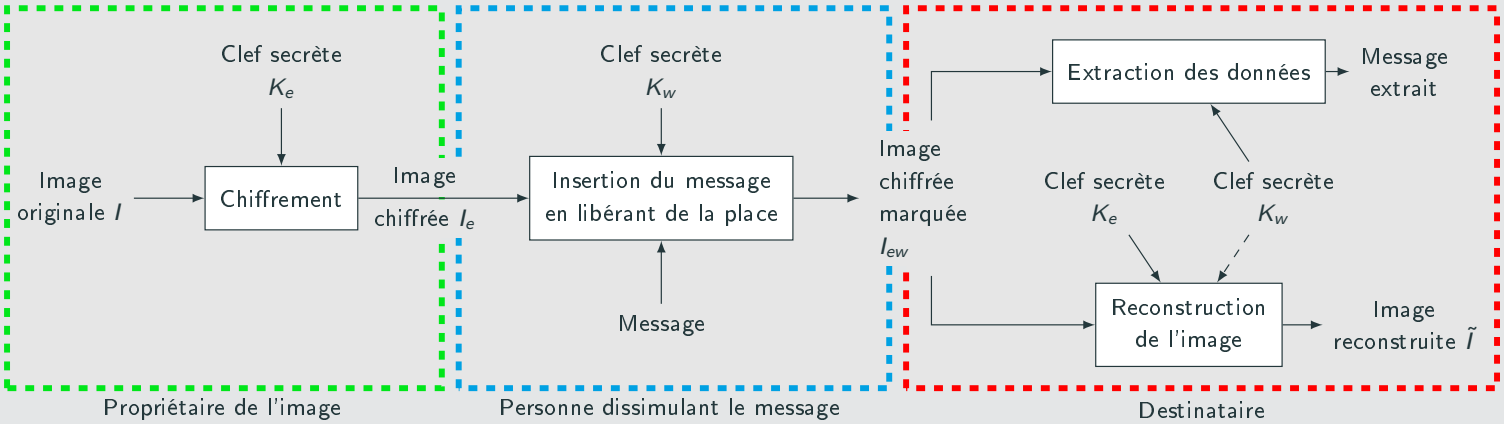
Sommaire

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



Etat de l'art

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

Etat de l'art

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 et un bloc 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

Etat de l'art

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 (≈ 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

Etat de l'art

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 (≈ 1 *bpp*), reconstruction exacte de l'image originale (avec la clef de chiffrement seulement)

Inconvénient

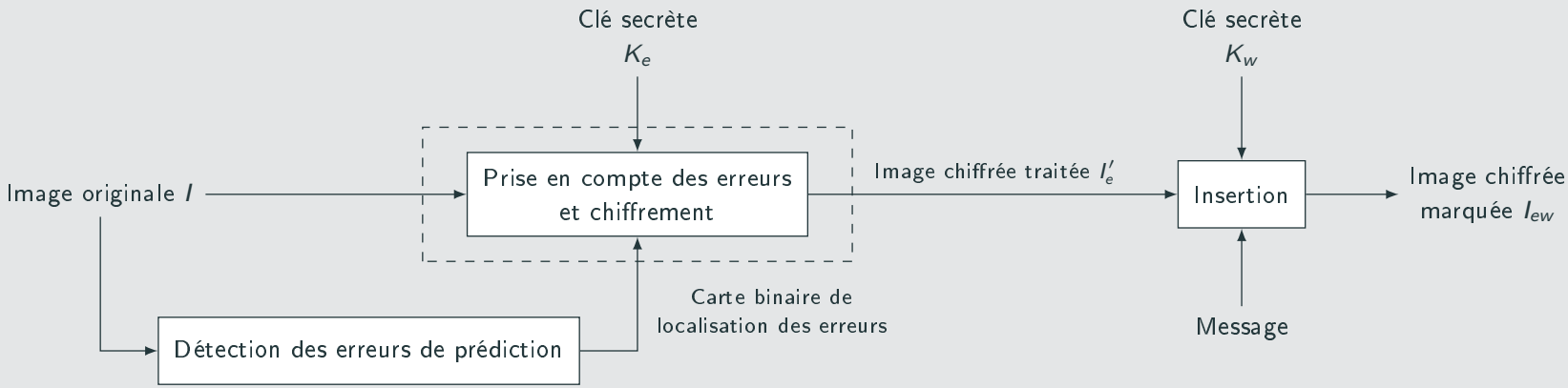
Données insérées non conservées après le déchiffrement de l'image

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

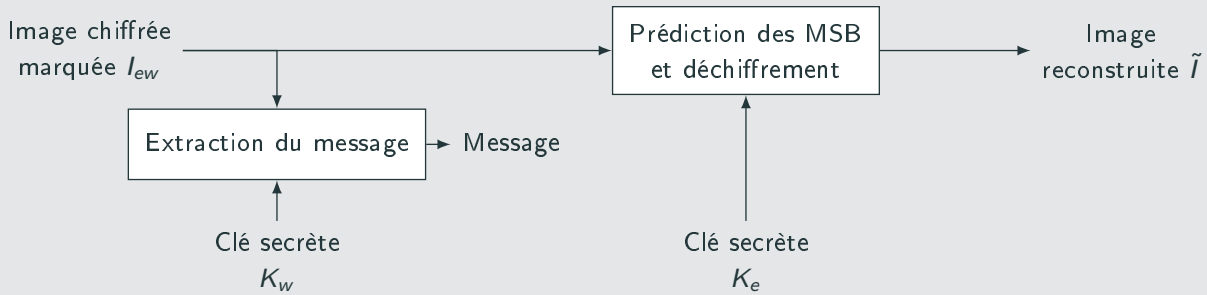
Sommaire

Schéma général

Chiffrement de l'image et insertion du message



Reconstruction de l'image et extraction du message



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 des MSB

Détection des erreurs de prédiction

- Pixel $p(i,j)$, d'inverse $inv(i,j) = (p(i,j) + 128) \mod 256$
- Prédicteur $pred(i,j)$ (d'après ses voisins déjà reconstruits)
  - Moyenne du voisinage
  - Pixel qui minimise la distance avec $pred(i,j)$
  - Prédicteur MED (JPEG-LS)
- Erreur de prédiction si $|pred(i,j) - p(i,j)| > |pred(i,j) - inv(i,j)|$

160	90
110	40

$p(i,j) = 40$, $inv(i,j) = (40 + 128) \mod 256 = 168$

Par exemple : $pred(i,j) = (90 + 110 + 160)/3 = 120$

$|pred(i,j) - p(i,j)| = |120 - 40| = 80$

$|pred(i,j) - inv(i,j)| = |120 - 168| = 48$

Erreur de prédiction (car $80 > 48$)

- Construction d'une carte de localisation des erreurs de prédiction

Deux approches proposées

Correction des erreurs de prédiction

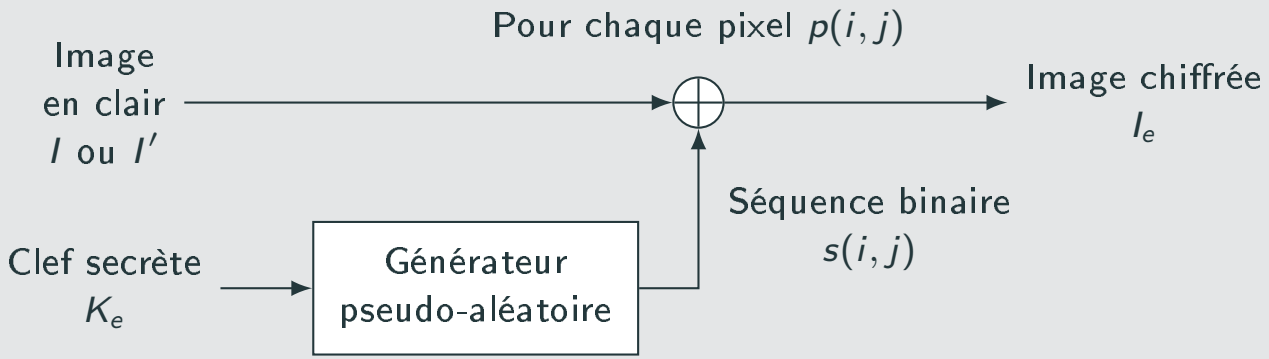
- Pré-traitement de l'image originale (adaptation des pixels pour éliminer les erreurs de prédiction)
- Chiffrement de l'image pré-traitée
- Insertion de données dans tous les pixels

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

Chiffrement et insertion du message

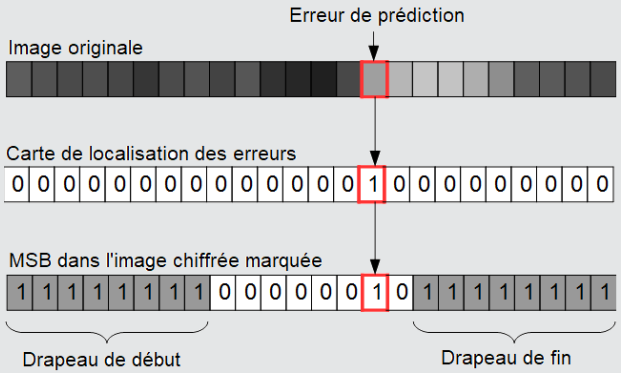
Chiffrement



Substitution des valeurs des MSB

$$p_{ew}(i,j) = b_k \times 128 + (p_e(i,j) \mod 128), \text{ pour } 0 \leq k < m \times n$$

- Approche_{1bpp} : pour tous les pixels
- $\text{Approche}_{réversible}$: pour insérer les drapeaux, puis le message



Extraction du message et reconstruction de l'image

Extraction du message secret

$$b_k = p_{ew}(i,j)/128, \text{ pour } 0 \leq k < m \times n$$

Reconstruction de l'image

- Approche$_{1bpp}$:
  - Aucune erreur de prédiction, donc déchiffrement en aveugle
  - Image pré-traitée parfaitement reconstruite
- Approche$_{réversible}$:
  - Prédiction des MSB (normale ou inverse), suivant les drapeaux
  - Totalement réversible, image originale parfaitement reconstruite

Résultats

Mesure des performances

- Erreurs lors de l'extraction du message
- Capacité d'insertion (*charge utile* en *bpp*)
- Qualité de l'image reconstruite (PSNR en *dB* et SSIM)

Pour chaque approche

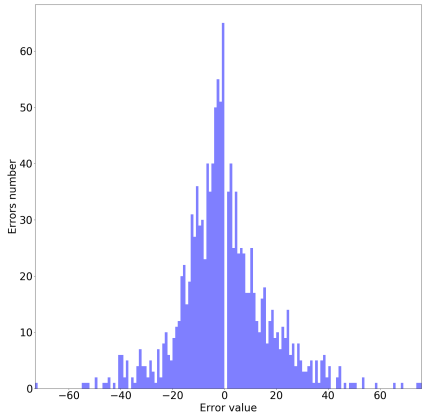
- Exemple détaillé
- Résultats sur la base BOWS-2 (10000 images PGM)
- Comparaisons avec l'état de l'art

Image originale.

Résultats

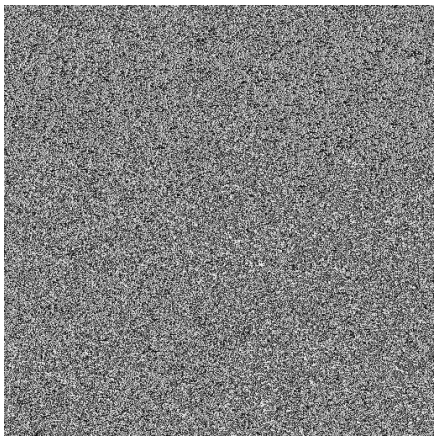
a)

b)

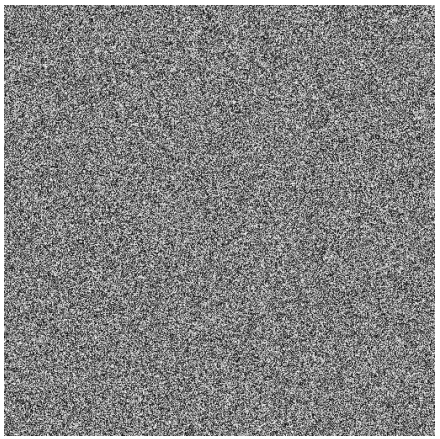


c)

d)



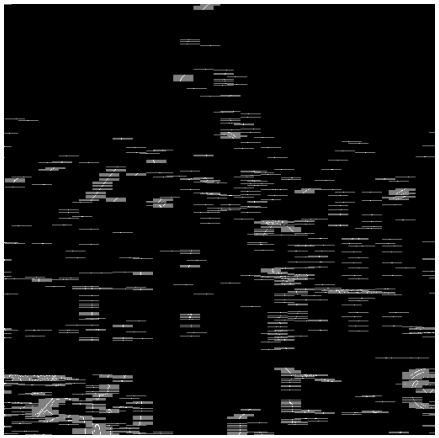
e)

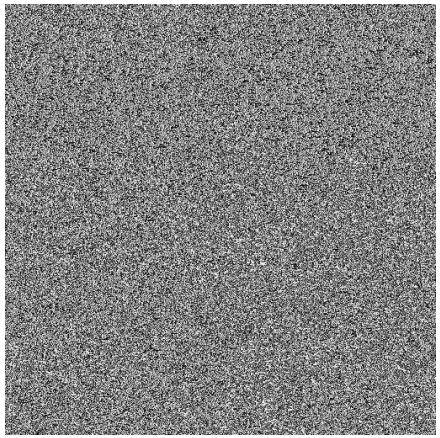


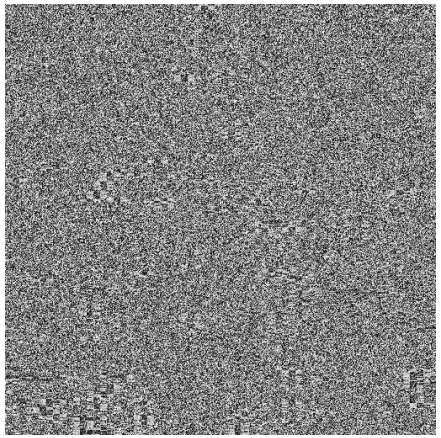
f)

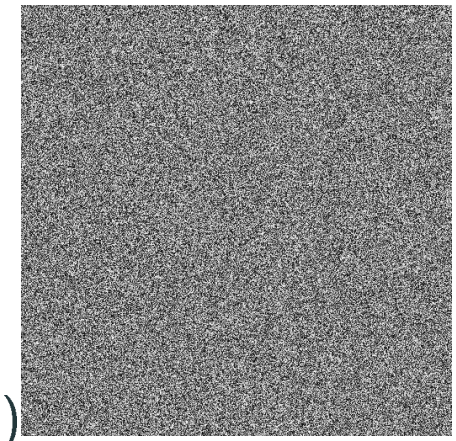
Exemple détaillé pour l'approche$_{1\,bpp}$: a) Emplacement des erreurs de prédiction, nombre d'erreurs = 1242 (0,47%), b) Histogramme des erreurs de prédiction, c) Image pré-traitée, PSNR = 46,87 *dB*, d) Image chiffrée, e) Image chiffrée marquée, capacité d'insertion = 1 *bpp*, f) Image reconstruite, PSNR = 46,87 *dB*, SSIM = 0,99.

Résultats

a) 

b) 

c) 

d) 

e) 

Exemple détaillé pour l'approche$_{\textit{réversible}}$: a) Emplacement des pixels non marqués (erreurs/drapeaux) nombre d'erreurs = 1225 (0,46%), b) Image chiffrée, c) Image chiffrée après signalement des erreurs de prédiction, d) Image chiffrée marquée, capacité d'insertion = 0,92 *bpp*, e) Image reconstruite, PSNR $\rightarrow +\infty$, SSIM = 1.

Résultats

		Meilleur cas $(6,3\%)$	Pire cas	Moyenne
Approche$_{1\,bpp}$	% d'erreurs de prédiction dans l'image originale	0%	$4,9\%$	$0,2\%$
	Capacité d'insertion (bpp)	1	1	1
	PSNR (dB)	$+\infty$	$29,0$	$57,4$
	SSIM	1	$0,98$	$0,99$
Approche$_{réversible}$	% d'erreurs de prédiction dans l'image originale	0%	$5,3\%$	$0,2\%$
	Capacité d'insertion (bpp)	1	$0,38$	$0,98$
	PSNR (dB)	$+\infty$	$+\infty$	$+\infty$
	SSIM	1	1	1

Performances de nos deux approches sur la base BOWS-2 (10000 images).

Résultats

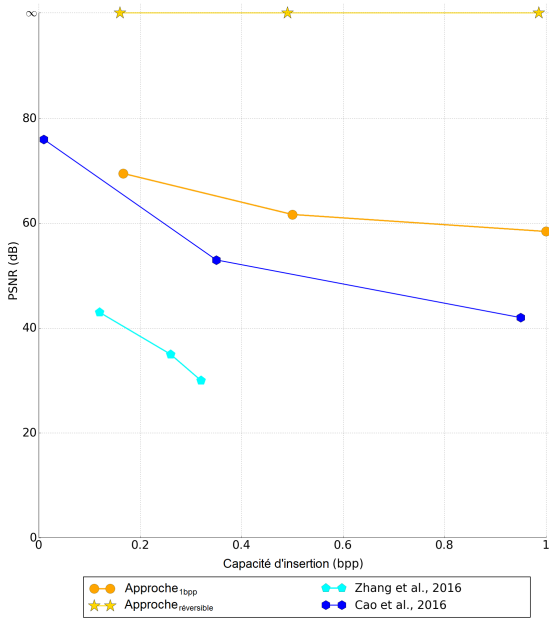


Image de test : Crowd.

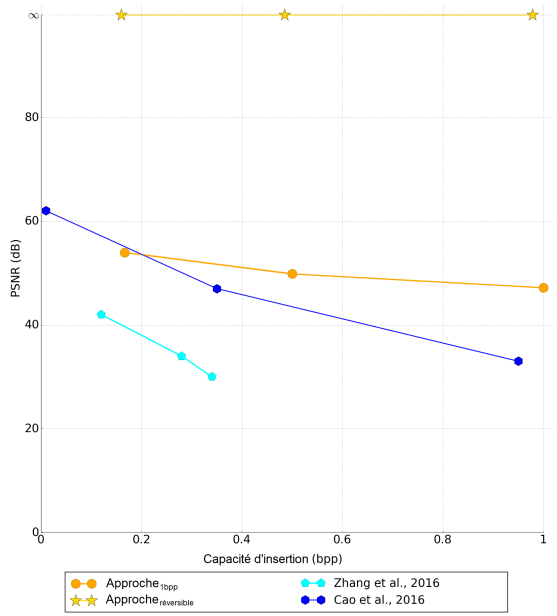


Image de test : Man.

Comparaisons de nos approches avec deux méthodes récentes de l'état de l'art.

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

X. Cao, L. Du, X. Wei, D. Meng, and X. Guo

High capacity reversible data hiding in encrypted images by patch-level sparse representation

IEEE Transactions on Cybernetics, vol. 46, no. 5, pp. 1132–1143, 2016

Sommaire

Conclusion

Résumé

- Approche$_{1\,bpp}$:
  - Correction des erreurs de prédiction
  - Insertion très haute capacité (capacité d'insertion $= 1$ bpp)
  - Bonne qualité de l'image reconstruite (PSNR $\approx 57,4$ dB)
- Approche$_{réversible}$:
  - Signalement des erreurs de prédiction
  - Haute capacité de dissimulation (capacité d'insertion $\approx 0,98$ bpp)
  - Image parfaitement reconstruite (PSNR $\rightarrow +\infty$)

Conclusion

Extensions

- Extension de l'approche$_{réversible}$ au deuxième MSB
  - Totalement réversible
  - Pour le deuxième MSB : capacité d'insertion $\approx 0,77$ *bpp*
  - Au total : capacité d'insertion $\approx 1,75$ *bpp*
- Méthode hybride, basée "poupées russes" (processus récursif)
  - Sur l'ensemble des plans binaires, du MSB au LSB (si possible)
  - Totalement réversible
  - Capacité d'insertion $\approx 2,56$ *bpp* en moyenne ($6,21$ *bpp* au max)
- Stéganalyse/cryptanalyse des images chiffrées marquées

Merci pour votre attention !

Questions ?