

Sujet de Thèse :

Stéganographie et stéganalyse d'images couleurs

L'objet de cette thèse portera à la fois sur l'insertion de messages indétectables (c'est le principe de la stéganographie, voir <https://fr.wikipedia.org/wiki/Stéganographie>) et sur la détection automatique de ceux-ci par l'analyse des statistiques des contenus observés (c'est le principe de la stéganalyse). Ces deux disciplines ont des méthodologies qui peuvent s'appliquer sur plusieurs types de contenus : la stéganographie cherchera à insérer le maximum d'information tout en minimisant la distorsion statistique du contenu après insertion) et la stéganalyse au contraire cherchera à détecter la présence d'un message caché en extrayant des caractéristiques qui sont statistiquement sensibles à l'insertion. Les contenus ici étudiés seront à savoir des images compressées couleurs, des contenus largement utilisées en pratique sur le web mais très peu étudiés par la communauté [3]. Les schémas de compressions privilégiés seront les standards JPEG et JPEG2000.

Les méthodes actuelles de stéganographie travaillent principalement sur des images en niveaux de gris ce qui empêche de prendre en compte la dépendance entre les 3 canaux couleurs. De récents travaux ont cependant montrés que l'utilisation de l'image brute (format RAW) pour générer l'image stégo (i.e. contenant un message caché) compressée permet de diminuer sensiblement la détectabilité des images [1]. Une stratégie similaire pourra être appliquée lors de l'étape du dématricage (voir voir <https://fr.wikipedia.org/wiki/Dématricage>). La distribution statistique du bruit d'acquisition du capteur sera également modélisée dans les domaines pixéliques, fréquentiels et colorimétriques, elle permettra de contraindre la distribution du signal inséré par l'algorithme de stéganographie. Enfin, la synchronisation spatiale et colorimétrique du canal d'insertion devra être mise en oeuvre afin d'avoir des modifications qui soient cohérentes dans un voisinage spatio-colorimétrique donné.

Les méthodes développées seront ensuite évaluées via la conception et l'utilisation de méthodes de stéganalyse dédiées aux images couleurs. Plusieurs pistes seront envisagées, à savoir l'utilisation d'une information adjacente liée aux probabilités d'insertion [2] pour compléter les caractéristiques de co-occurrences [4], l'utilisation des dépendances fortes qu'entraîne le dématricage couleur, ou encore l'extraction supervisée de caractéristiques via l'utilisation de réseaux de neurones profonds [5].

Connaissances et thèmes abordés : modélisation du signal et des images - apprentissage statistique

Capacités en programmation : Python - Matlab

Début de la thèse : Septembre 2016

Contact :	Patrick Bas - Directeur de Recherche au CNRS
Email :	patrick.bas@ec-lille.fr
Adresse :	CRIStAL, UMR CNRS 9189 Ecole Centrale de Lille - Bureau C315

Références

- [1] T. DENEMARK ET J. FRIDRICH, *Side-informed steganography with additive distortion*, in IEEE International Workshop on Information Forensics and Security, 2015.
- [2] T. DENEMARK, V. SEDIGHI, V. HOLUB, R. COGRANNE ET J. FRIDRICH, *Selection-channel-aware rich model for steganalysis of digital images*, in IEEE Workshop on Information Forensic and Security, Atlanta, GA, 2014.
- [3] A. D. KER, P. BAS, R. BÖHME, R. COGRANNE, S. CRAVER, T. FILLER, J. FRIDRICH ET T. PEVNÝ, *Moving steganography and steganalysis from the laboratory into the real world*, in Proceedings of the first ACM workshop on Information hiding and multimedia security, ACM, 2013, p. 45–58.
- [4] T. PEVNÝ, P. BAS ET J. FRIDRICH, *Steganalysis by subtractive pixel adjacency matrix*, Information Forensics and Security, IEEE Transactions on, 5 (2010), p. 215–224.
- [5] L. PIBRE, P. JÉRÔME, D. IENCO ET M. CHAUMONT, *Deep learning for steganalysis is better than a rich model with an ensemble classifier, and is natively robust to the cover source-mismatch*, arXiv preprint arXiv :1511.04855, (2015).