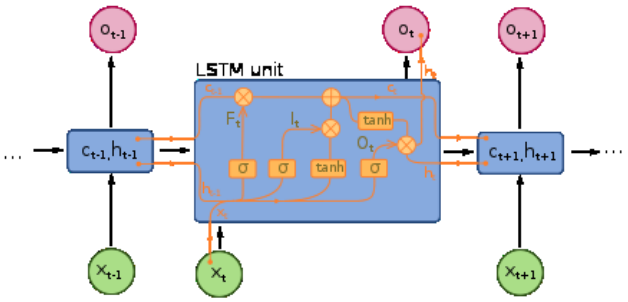

Stage : " 'Pooled steganalysis' par Deep-Learning "

Marc CHAUMONT, Ahmad ZAKARIA, Gérard SUBSOL, Hugo RUIZ
LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier)
Equipe ICAR, 860 rue de St Priest, 34095 Montpellier cedex 5- France
Tel : +33 4.67.14.97.59, Marc.Chaumont@lirmm.fr

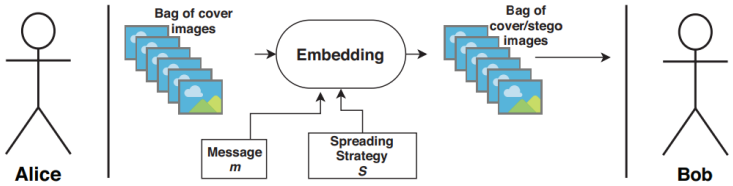

Fig. 1. A scheme that illustrates how the steganographer, Alice, spreads a message $\mathbf{m}$ in multiple covers using a strategy $s \in \mathcal{S}$.

Mots clefs : Traitement d'images, Stéganographie, Stéganalyse, Machine-Learning, Deep-Learning.

La stéganographie / stéganalyse peut être expliquée comme un jeu à trois participants. Les stéganographes classiquement appelés Alice et Bob, souhaitent envoyer un message, dont l'existence même n'est connue que d'eux seul. La stéganalyste, généralement appelé Eve, observe les échanges d'images qui ont lieu entre Alice et Bob et cherche à déterminer si Alice et Bob communiquent [Simmons83]. La stéganographie est donc l'art de dissimuler un message dans un support, ici une image, pour le transmettre de manière secrète, et la stéganalyse est l'art de déceler la présence de ce message ou non dans le support. Cette discipline dans sa version moderne, c'est-à-dire numérique, a débuté au début des années 2000.

Nous nous intéressons ici à la "*batch steganography*" (stéganographie par lots) et la "*pooled steganalysis*" (stéganalyse groupée). Dans le cadre de la "*batch steganography*", Alice dissimule (dispatche) un message dans un ensemble de supports (c.à.d. un ensemble d'images). La Fig. 1 schématise ce principe. De son côté, Eve doit décider si un ensemble d'images (un "*bag*") dissimule un message. Eve doit donc accumuler un ensemble de preuves pour décider, si oui ou non, Alice et Bob communiquent.

Dans les récents travaux menés au sein de l'équipe ICAR, par Ahmad Zakaria (publication en **décembre 2019** [Zakaria et al. 2019]), nous avons expérimenté de nombreuses stratégies d'insertions avec divers scénarios de steganalyse. Nous avons préparé plusieurs bases d'images, et utilisé une architecture moderne adaptée à la "*pooled steganalysis*". L'architecture par machine-learning (mais ce n'est pas du deep learning) est schématisé en Fig. 3. L'architecture, bien que générique, permet d'obtenir de bonnes performances. Notre protocole d'apprentissage permet quant à lui d'augmenter les performances de "*pooled steganalysis*".

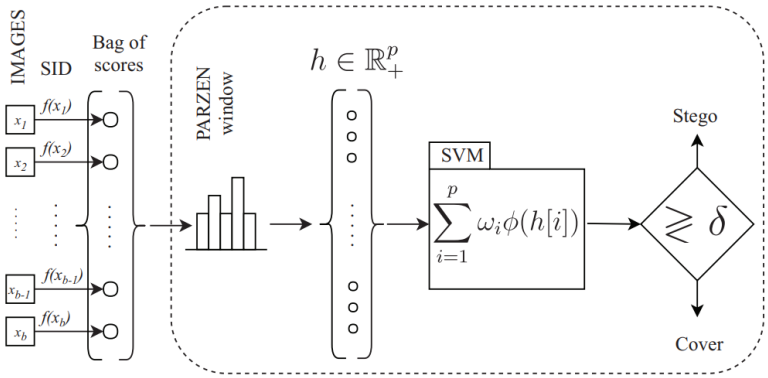

Fig. 3. The general pooled steganalysis architecture from [6]. ϕ is the re-description transformation function. δ is a threshold.

L'objectif du stage est de s'inspirer de l'architecture de "*pooled steganalysis*" que nous avons utilisé dans [Zakaria et al. 2019], et de la **basculer en une architecture de type deep learning**. De nombreuses questions restent ouvertes, notamment la gestion de l'invariance à l'ordre des images, la gestion d'images de différentes tailles, la gestion d'un apprentissage multi-classes, et éventuellement la robustesse au nombre de bits insérés. Cela dit, il existe plusieurs options, comme par exemple effectuer un apprentissage utilisant des "*Convolutional Neural Network*" (CNN), ou alors utiliser des réseaux récurrents (LSTM).

Pour mener à bien ce sujet, il est préférable d'avoir certaines connaissances : en traitement des images, et/ou en classification/fouille de données, et/ou en architecture des machines/installation d'OS. Il est également intéressant d'avoir de bonnes bases en programmation, dont Python, et en mathématiques. Une connaissance en Deep Learning est un plus.

Profil recherché : Master (M2) ou Ecole d'Ingénieur (3ème année) ayant une bonne maîtrise de la programmation (C++, Python…), des connaissances en fouille de données / indexation / classification, traitement des images, sécurité.

Encadrement : Marc CHAUMONT (Enseignant Chercheur), Gérard SUBSOL (Chercheur), Ahmad ZAKARIA (Doctorant), Hugo RUIZ (Ingénieur d'étude).

Modalité de candidature : Envoyez un CV, une lettre de motivation ainsi que votre relevé de notes de M1 le plus tôt possible. Après pré-sélection des candidatures, des entretiens téléphoniques ou en personne seront planifiés.

Contacts : Marc Chaumont (*marc.chaumont@lirmm.fr*)
Lieu du stage : LIRMM, équipe ICAR.
Période du stage : 1er semestre 2020 (5-6 mois).
Gratification de stage : environ 550€ mois.

Bibliographie:

[Simmons83] G. J. Simmons, "The prisoners problem and the subliminal channel," in Advances in Cryptography, CRYPTO, Aug. 1983, pp. 51–67.

[Zakaria et al. 2019] Ahmad Zakaria, Marc Chaumont, Gerard Subsol, " Pooled Steganalysis in JPEG: how to deal with the spreading strategy? ", **WIFS'2019**, IEEE International Workshop on Information Forensics and Security, December 9-12, 2019, Delft, The Netherlands, 6 pages, Acceptance rate = 30%. ArXiv version http://arxiv.org/abs/1906.11525.