Nom et Prénom : KICH ISMAIL
Date de soutenance : 09/10/2021
Directeur de Thèse : AMEUR EL BACHIR

Sujet de Thèse :

Modèles de dissimulation d'information par Machine Learning

Résumé :

La stéganographie est la technique de dissimulation de l'information secrète dans des supports numériques anodins d'une manière invisible. Elle réalise la communication secrète, à l'inverse de la cryptographie qui sert à crypter l'information lors sa transmission dans un canal de communication publique. L'objectif de la stéganographie est que personne d'autre que le destinataire ne doit suspecter l'existence de l'information secrète au sien du support hôte transféré. Plusieurs techniques et approches ont été proposées. Parmi ces approches, nous nous sommes intéressés à celles basées sur les algorithmes d'apprentissage automatique. Dans cette thèse, nous avons d'abord présenté les notions relatives à la stéganographie et à l'apprentissage automatique.

Ensuite, à l'aide de l'apprentissage non supervisé, nous avons développé deux modèles stéganographiques basés sur le clustering dont l'objectif est la dissimulation d'un message texte dans une image de couverture, le premier modèle est basé sur l'algorithme de clustering k-means, tandis que dans le deuxième modèle nous avons utilisé l'algorithme Modified Simple Linear Iterative Clustering (M-SLIC) et la méthode de correction $2r$ afin de minimiser la distorsion de l'image de couverture. Par la suite, en utilisant l'apprentissage profond, nous avons développé deux autres modèles dont l'objectif est de dissimuler une image couleur dans une autre image couleur de même taille. Le premier modèle est un schéma stéganographique basé sur un réseau auto-encodeur à convolutions classiques, le deuxième modèle est basé sur un réseau autoencodeur à convolutions Dilatées. Nous avons développé pour le deuxième modèle une nouvelle fonction de perte qui favorise la qualité de l'image stégo durant l'entrainement. Nous avons implémenté et testé ces modèles sur différentes bases de données d'images numériques.

Mots clés : Sécurité d'information, Stéganographie, Apprentissage automatique, Clustering, Réseaux de neurones convolutionnels, Auto-Encodeur, Apprentissage profond

Absract :

Steganography is the technique of concealing invisibly a secret information into innocent looking digital files. It achieves secret communication unlike cryptography that encodes the information before transmitting it into a public communication channel. The objective of steganography is that no one other than the recipient should suspect the existence of secret information inside the transmitted host file. Several techniques and approaches have been proposed. Among these approaches, we were interested into the ones based on machine learning. In this thesis, we firstly presented the different notions related to steganography and machine learning. Then, using unsupervised machine learning, we developed two steganographic models based on clustering that dissimulate text into cover image. The first model is based on K-means clustering, the second one is based on Modified Simple Linear Iterative Clustering (M-SLIC) and the $2r$ correction in order to minimize distortion on cover image. After that, using deep learning, we proposed two models aiming to embed a color image into another color image with the same size. The first one is based on auto-encoder with classical convolutions, the second one is based on auto-encoder with dilated convolutions. We developed for the second model a new loss function that favorites the quality of the stego image during training. We implemented and tested these models on digital images from different databases.

Keywords: Information Security, Steganography, Machine learning, Clustering, Convolutional Neural Network, Auto-Encoder, Deep Learning