Carnet de recherche

par Ayouba Anrezki
Initialisation : 12/03/2024
MÀJ : 12/03/2024

1 : 12/03/2024 : Définition du sujet.

1.1 : Définition du problème

La stéganographie désigne l'art de dissimuler de l'information de manière subtile. Toute la sécurité de cette méthode de dissimulation réside dans la non-connaissance des observateurs non avertis, de la présence d'une information cachée. La variante informatique de ce procédé consiste dans la dissimulation des données dans le corps d'autres données. Si la stéganographie permet de transférer des données à l'abri des regards des observateurs non avertis, nous pouvons toujours nous demander si il n'est pas possible d'affaiblir la sécurité de cette méthode de dissimulation. **Autrement dit, est-il possible de distinguer le bruit d'une information cachée ?**

1.2 : Idée d'orientation

Il existe un champ de recherche à part entière qui s'intéresse à la distinction entre données pures et données issues d'un processus stéganographique qui se nomme Stéganalyse

1.2.1 : Méthodes de distinction

- **Analyse statistique :**
  Les données qui contiennent simplement du bruit peuvent avoir des caractéristiques statistiques différentes de celles qui cachent des informations. Vous pourriez étudier des mesures telles que l'entropie, la distribution des valeurs de pixels, les corrélations spatiales, etc.
- **Analyse de la fréquence :**
  Les images qui cachent d'autres images peuvent avoir des motifs de fréquence différents de ceux des images contenant seulement du bruit. Les techniques de transformée de Fourier ou d'ondelettes peuvent être utiles pour analyser ces différences.
- **Analyse visuelle :**
  Même si les données semblent similaires visuellement, il peut y avoir des artefacts ou des modèles non perceptibles à l'œil nu. Vous pourriez explorer des techniques de traitement d'image avancées pour mettre en évidence ces différences.
- **Apprentissage automatique :**
  Vous pourriez également explorer des approches basées sur l'apprentissage automatique, où vous entraînez un modèle à différencier les deux types de données à partir d'un ensemble d'exemples étiquetés.

2 : L'analyse statistique

3 : 02/04/2024 : Phénomène aléatoire

3.1 : Entropie de Shannon

3.2 : Théorie de l'information

4 : 02/04/2024 : Mesure

5 : 12/09/2024 - Implémentation Ocaml

- Implémentation Ocaml des algorithmes pour la stéganographie image et

6 : 19/09/2024 - git init et recherche documentaire.

6.1 : Définition de la problématique

Problématique : Est-il possible de créer un algorithme de stéganalyse généraliste, i.e. un algorithme qui n'a pas connaissance du mode de dissimulation utilisé ?

Les différents modes de dissimulation :

- Système de substitution : remplacer une partie de la cover (1) par des données de l'information à dissimuler.
- Transformation des paramètres de la cover : modification des paramètres physiques de la cover en fonction de l'information à dissimuler (ex : fréquence).
- Même chose avec le spectre.
- Méthode statistique : modifier la distribution statistique de la cover en fonction de la stégo.
- Techniques de distorsion : stocker des informations par distorsion du signal et mesurer l'écart par rapport à la couverture originale lors de l'étape de décodage.
- Méthodes de génération de couverture : encoder les informations de manière à cacher un secret dans la communication créée.

Objectif : Trouver un invariant de dissimulation !

7 : 26/09/2024 : Prolongement par continuité de la semaine dernière (lecture 10)

- **Problématique :** Est-il possible d'identifier un pattern, une caractéristique propre aux données issues du processus de stéganographie ?

7.1 : Protocole :

- Étudier les différentes méthodes de stéganographie (substitution, transformation, spectre, statistique, distorsion, génération de cover).
- Étudier la réponse stéganalyse à ces algorithmes.
- Identification d'invariants de dissimulation.

7.2 : 03/10/2024 : Définition formelle de l'information :

[
l1 |(1,0,1) (1,0,1)|
l2 |(0,0,0) (0,0,0)|
]

Une information est une matrice de tuples de taille n de nombres binaires.

- **Cas de base :**
  - **Information vide (null) :**

    On note ε l'information vide de taille $|\varepsilon| = 0$
    $\varepsilon = ()$

  - **Information de base :**

    $\forall\ (b_n) \in \mathbb{B}^{\mathbb{N}}$ fini $L = \left((b_0 b_1 ... b_n)\right)$ de taille $|L| = n + 1$

  - **Notation**

- On note $\mathcal{M}_{n,p,l}(\mathbb{B}^{\mathbb{N}})$ l'ensemble des informations de matrice dans $\mathcal{M}_{n,p}(\mathbb{B}^{\mathbb{N}})$ dont les tuples sont de l éléments.

- **Opérations sur les informations :**
  - **Taille d'une information :** Soit $L \in \mathcal{M}_{n,p}(\mathbb{B}^{\mathbb{N}})$ une information, la taille de L est notée $|L| = n \times p$
  - **Caractéristiques d'une information :** Soit $L \in \mathcal{M}_{n,p}(\mathbb{B}^{\mathbb{N}})$ une information
  - **Union/Intersection :** Soit L_1 et L_2 deux informations de taille n
    - $L_1 \cup L_2 =$

8 : 10/10/2024 : Définition du répertoire documentation/ prototypage

8.1 : Définition formelle de l'information

9 : Vocabulaire (MAJ 12/03/2024)

1. données pures : données ne cachant pas d'autres données issues d'un processus stéganographique.
2. cover : support pour la dissimulation d'informations cachées.
3. stego : information à cacher.
4.

10 : Lecture en attente :

1. https://utt.hal.science/hal-02470070/document
2. https://fr.wikipedia.org/wiki/Entropie_de_Shannon
3. https://fr.wikipedia.org/wiki/Th%C3%A9orie_de_l%27information
4. https://hal.science/hal-00394108/document
5. https://greenteapress.com/thinkdsp/thinkdsp.pdf
6. http://tinyurl.com/thinkdsp08
7. https://fr.wikipedia.org/wiki/Algorithme_de_Knuth-Morris-Pratt
8. https://theses.hal.science/tel-00706171v2/file/RCogranne_soutenance.pdf
9. https://repository.root-me.org/St%C3%A9ganographie/FR%20-%20Analyse%20st%C3%A9ganographique%20d%27images%20num%C3%A9riques.pdf
10. https://d1wqtxts1xzle7.cloudfront.net/11025045/22359536_lese_1-libre.pdf?1363619886=&response-content-disposition=inline%3B+filename%3DA_survey_of_steganographic_techniques.pdf&Expires=1726758425&Signature=UWNEvv4JIxHsL-iZcX-PzwvRlbmce0~unnnAUFS2lB~tsuJUbrH1Mzt4ZnO~D1Dhn9DKUo0jtG-BZnkuZYYz5iSvTUuJHJJqcZ65yceho5qgmi7Jpv9OnJsNLxnqAjhHp~frVhRI3yYvhmZRsOL0gdCCCy6O5Bb9XcylGQC8DG2cJBk-1GRz5NbPu5Udq4R1U-pr2GvYZKJJmqnb7MQoutftG~9-jS~WMxnag3IlAe8g~vlz87mWWLxGle-6fbBg1I-EOa63b3fzUVsFY2bLQo0WgwqNMQ__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA