UNIVERSITÉ DE YAOUNDÉ I
***
ÉCOLE NATIONALE SUPÉRIEURE POLYTECHNIQUE DE YAOUNDÉ
***
DÉPARTEMENT DE GÉNIE INFORMATIQUE
***
HUMANITÉS NUMÉRIQUES

THE UNIVERSITY OF YAOUNDE I
***
NATIONAL ADVANCED SCHOOL OF ENGINEERING OF YAOUNDE
***
DEPARTMENT OF COMPUTER ENGINEERING
***
DIGITAL HUMANITIES

Exercices Livre 2

Rédigé par :

FANTA YADON Félicité 22P069

Supervisé par :

M. Thierry MINKA

I Partie 1 : Fondements Philosophiques et Épistémologiques

Exercice 1 : Analyse Critique du Paradoxe de la Transparence

Le paradoxe de la transparence, tel que conceptualisé par Byung-Chul Han, stipule que l'impératif sociétal et administratif de la transparence totale, loin de créer une société plus honnête, aboutit à une aliénation par l'hyper-visibilité et fragilise la sphère intime. En investigation numérique, cela se traduit par une tension aiguë entre la nécessité d'accès aux données pour l'établissement de la vérité et le droit fondamental à la vie privée de l'individu.

II Partie 1 : Fondements Philosophiques et Épistémologiques

II.1 Exercice 1 : Analyse Critique du Paradoxe de la Transparence

L'investigation numérique, confrontée à l'impératif éthique et légal de la transparence, se heurte inévitablement au « Paradoxe de la Transparence » popularisé par Byung-Chul Han, où l'excès de visibilité engendre une forme nouvelle et plus subtile d'aliénation et de vulnérabilité. L'enjeu est de concilier la recherche de la vérité factuelle avec le respect de la dignité humaine.

1. **Introduction (10%)** Le paradoxe est introduit : la vérité numérique exige la *traçabilité* intégrale des actions et des communications (impératif de preuve), mais cette traçabilité illimitée contredit directement les droits fondamentaux tels que le droit à l'oubli et le respect de la dignité. Cette tension est juridiquement incarnée en Droit français par l'opposition entre l'Article 15 de la Déclaration des Droits de l'Homme et du Citoyen (DDHC, pour la transparence gouvernementale) et l'Article 9 du Code Civil (pour le respect de la vie privée).

2. **Développement (70%)**

    a. **Analyse du Cas Concret** : Considérons une enquête pour fuite de données (Affaire X) où un investigateur obtient un accès total aux journaux de connexion, aux métadonnées et aux communications d'un suspect potentiel. Bien que cet accès permette de réunir la *vérité factuelle* relative à l'acte, il contamine inéluctablement l'enquête par un **excès d'informations privées** non pertinentes (opinions politiques, relations personnelles, données de santé). Cet accès illimité et non ciblé risque de violer la dignité de la personne, traitée comme un simple amas de données, avant même que sa culpabilité ne soit établie.

    b. **Application de l'Éthique Kantienne** : L'impératif catégorique de Kant exige que l'action (*maxime*) soit universalisable sans contradiction logique. Si la maximisation de la transparence des systèmes mène systématiquement à la destruction de l'intimité, elle ne peut être érigée en loi morale universelle. L'investigateur doit donc opérer selon une maxime éthique plus stricte :

> « Maximiser la vérité factuelle nécessaire à la justice, tout en préservant le statut de la personne comme **fin en soi**, et non comme simple moyen d'accéder aux données. »

c. **Proposition de Résolution Pratique et Réaliste** : La résolution pratique réside dans les technologies cryptographiques qui découplent l'accès à la donnée de la vérification de sa validité. La **Preuve à Divulgation Nulle de Connaissance avec Non-Répudiation (ZK-NR)** est un exemple de solution idéale. Cette technologie permet à l'investigateur de **vérifier** l'authenticité d'une trace numérique ou l'existence d'une action spécifique (**Vérification sans Divulgation**), sans jamais avoir à accéder aux données brutes. Cette approche concilie l'établissement rigoureux de la vérité légale avec le respect absolu de la confidentialité et de la dignité.

3. **Conclusion (20%)** La recherche de la vérité en investigation numérique ne doit pas être le simple produit d'un accès illimité, mais le résultat d'une **méthodologie éthique et technologique**. Le ZK-NR et les techniques similaires représentent une **Innovation Épistémologique** fondamentale, car ils déplacent le lieu de la vérité de la *divulgation* totale à la *vérification probabiliste* et cryptographique. L'avenir de l'investigation est dans la **vérification sans surveillance**.

III Partie 1 : Fondements Philosophiques et Épistémologiques

III.1 Exercice 2 : Transformation Ontologique du Numérique

La transition vers la société numérique a engendré une mutation des fondements de l'existence humaine (*l'être*) et de la preuve légale. Ce phénomène est analysé ici sous l'angle de l'ontologie heideggérienne et de la phénoménologie de la trace numérique.

III.1.1 2.1 Comparaison Heidegger / Numérique

- **Heidegger** : Le concept de *Dasein* (l'« être-là ») est central. L'existence est ancrée dans son **être-au-monde** physique, caractérisée par la temporalité et la finitude (*Être-pour-la-mort*). L'essence de l'homme se révèle dans son rapport aux outils (*l'étant à portée de main*).
- **Numérique** : La trace numérique crée le *Dasein Augmenté* ou *Double Numérique*. L'existence est étendue au-delà du corps physique. Ce double est constitué par les données, les interactions en réseau et l'archivage perpétuel (*digital persistence*). Le numérique est une nouvelle modalité ontologique où l'être peut **persister** ou **agir** sans être physiquement présent, défiant les notions classiques d'ici et de maintenant.

III.1.2 2.2 Analyse de Profil Social : Identification des Traces Exstentielles

Un profil social n'est pas une simple base de données, mais une **manifestation phénoménologique** de l'existence de l'individu. L'investigateur doit y lire les structures fondamentales de l'être au sens heideggérien à travers les cinq « traces existentielles » :

- **Relations Sociales (*Mitsein* / Être-avec)** : Les listes d'amis, les abonnements, les mentions et les partages. Ces traces cartographient les relations intersubjectives et l'appartenance à une communauté numérique, définissant l'« être-avec » dans l'espace digital.
- **Centres d'Intérêt (*Sorge* / Souci)** : Les pages aimées, les groupes suivis, les contenus commentés ou sauvegardés. Ces éléments révèlent l'objet de la préoccupation de l'être, son champ d'investissement et sa **projection de soi** dans le monde thématique.
- **Temporalité (*Historizität* / Historicité)** : La fréquence de connexion, les archives de messages, les dates de publication. Ces données définissent l'historicité et la durée du Dasein numérique. Elles permettent de reconstituer la **chaîne des événements** et l'évolution temporelle des intentions.
- **Spatialité (*Mondanité* / Être-au-monde)** : Les géolocalisations, les étiquettes de lieux, les enregistrements de voyage. Ces traces ancrent l'existence numérique dans un espace physique, établissant le lien critique entre le **double numérique** et l'être corporel.
- **Intentionalité (*Projection* / Entwurf)** : Les contenus créés, les manifestes politiques, les messages directs à caractère délibératif. Ces traces révèlent l'intention et la projection de l'être dans l'action, permettant de reconstruire les motivations profondes à l'origine de l'acte incriminé.

III.1.3 2.3 Impact sur la Preuve Légale

La preuve légale subit une transformation ontologique fondamentale :

- **De l'Objet au Processus** : La preuve n'est plus un « objet » statique (document papier, arme, témoignage) ayant une existence autonome. Elle devient un **« processus relationnel »** dynamique, constitué de flux, de logs et de métadonnées interconnectées.
- **Reconstruction de la Vérité** : La vérité n'est plus « trouvée » (par simple observation de l'objet), mais **reconstruite** par la corrélation et l'interprétation des traces ontologiques. L'investigateur ne cherche pas un fait isolé, mais la **structure relationnelle totale** qui a rendu l'événement possible.
- **Complexification de la Validation** : La validité de la preuve dépend désormais non seulement de son authenticité technique (non-altération), mais aussi de son **contexte interprétatif**. Son sens dépend de sa place dans la structure relationnelle totale du Dasein numérique, ce qui complexifie sa recevabilité et son poids devant les tribunaux.

IV Partie 2 : Mathématiques de l'Investigation

IV.1 Exercice 3 : Calcul d'Entropie de Shannon Appliquée

L'entropie de Shannon (H) est une mesure de l'incertitude ou de l'aléa dans une source d'information. En investigation numérique, un niveau d'entropie élevé (proche de 8 bits/octet) est un indicateur fort de la présence de données **chiffrées**, **compressées** ou de **malwares**.

IV.1.1 3.1 Formule de l'Entropie et Code de Correction

Formule de l'Entropie de Shannon pour une source discrète :

$$H = -\sum_{i} p(x_i) \log_2(p(x_i))$$

où $p(x_i)$ est la probabilité d'occurrence du caractère ou de l'octet x_i. La valeur maximale est $H_{max} = 8$ bits/octet.

Code de Correction Attendu (Python) :

Listing 1: Calcul de l'entropie de Shannon en bits/octet

```
import math
from collections import Counter

def calculate_entropy(data):
    """Calcule l'entropie de Shannon en bits/octet"""
    if len(data) == 0:
        return 0

    # 1. Comptage des fr quences des octets
    byte_counts = Counter(data)
    total_bytes = len(data)

    entropy = 0.0
    for count in byte_counts.values():
        # 2. Calcul de la probabilit  de chaque byte
        p_x = count / total_bytes
        # 3. Contribution    l'entropie
        # Utilisation du logarithme en base 2 (log2) pour obtenir le r sultat
            en bits
        entropy -= p_x * math.log2(p_x)

    return entropy

# Exemple de validation : Entropie d'un flux al atoire
# data_exemple = bytes([1, 5, 200, 4, 15] * 50)
# print(f"Entropie calcul e : {calculate_entropy(data_exemple):.4f} bits/octet
    ")
```

IV.1.2 3.2 Développement de l'Interprétation en Criminalistique Numérique

L'entropie n'est pas une preuve en soi, mais un puissant **indice de processus** qui permet de distinguer la **normalité statistique** de l'**anomalie intentionnelle**. L'analyse de l'entropie est cruciale en début d'investigation pour le triage des données (*triage analysis*).

IV.1.2.1 Critère d'Anomalie : Le Seuil Critique H ≈ 7.8 bits/octet Une valeur d'entropie dans la plage [7.5, 8.0] signale une distribution des octets **quasi uniforme**, typique d'un flux binaire parfaitement aléatoire (ou pseudo-aléatoire). Ce seuil est une **alerte rouge** pour l'investigateur, car la nature statistique des données textuelles, des bases de données ou des exécutables standards maintient leur entropie bien en deçà de 5 bits/octet.

Le dépassement de ce seuil critique impose la recherche immédiate de trois phénomènes :

1. **Détection de Chiffrement (Ransomware)**

   - **Principe** : Les algorithmes de chiffrement robustes (ex: AES) visent à garantir que chaque bit du texte clair a un impact maximal sur la distribution des octets du texte chiffré. Le résultat est un flux qui satisfait le test d'aléatoire, menant à $H \to 8$.
   - **Implication** : La transition soudaine de fichiers d'une entropie faible (ex: document Word à $H \approx 4.6$) à $H \approx 7.8$ est la signature numérique d'une **action de chiffrement malveillante** (ex: attaque par rançongiciel). L'investigateur confirme ainsi le *modus operandi* de l'attaque.

2. **Identification de Malwares Packés et Obfusqués**

   - **Principe** : Les auteurs de malwares utilisent des **packers** pour chiffrer ou compresser le code exécutable afin de contourner les moteurs antivirus basés sur les signatures statiques.
   - **Implication** : L'**Analyse d'Entropie Locale** d'un exécutable révèle des pics d'entropie (segments $H > 7$) au milieu d'une structure globale faible. Ce pic correspond à la zone du binaire qui contient la charge utile du malware **chiffrée**. Cette découverte indique la nécessité immédiate d'une ingénierie inverse dynamique (*reverse engineering*) pour dépacker le code en mémoire.

3. **Révélation de Stéganographie**

   - **Principe** : Bien que le masquage d'un message non chiffré puisse avoir un impact modéré, l'insertion d'un **message chiffré** (haute entropie) dans les bits de poids faible d'un fichier porteur (ex: une image) altère l'équilibre statistique.
   - **Implication** : La cartographie de l'entropie sur l'ensemble du fichier porteur peut révéler une **concentration anormale** de la haute entropie dans la zone du message caché, permettant à l'investigateur de cibler l'extraction.

L'entropie de Shannon est, au sens épistémologique, un outil qui permet à l'investigateur de transformer une **donnée brute ambiguë** en une **information classifiée et actionable**, fondement de toute prise de décision en gestion d'incident.

Interprétation du Résultat :

V Partie 2 : Mathématiques de l'Investigation (Détail de l'Entropie de Shannon)

V.1 Exercice 3 : Calcul et Interprétation de l'Entropie de Shannon Appliquée

Le concept d'Entropie de Shannon (H) est un pilier de la Théorie de l'Information. En criminalistique numérique, il est élevé au rang d'outil de diagnostic essentiel pour l'investigateur, permettant de **classifier rapidement la nature intrinsèque** d'un flux de données avant même de l'analyser sémantiquement.

V.1.1 3.1 Fondements et Formule

L'entropie mesure le degré d'**incertitude** ou d'**aléatoire** dans la distribution des octets d'un fichier. Pour une source discrète, elle est calculée par :

$$H = -\sum_{i} p(x_i) \log_2(p(x_i))$$

Où $p(x_i)$ est la probabilité d'occurrence de l'octet x_i. La valeur maximale pour un octet (8 bits) est $H_{max} = 8$ bits/octet, atteinte uniquement si tous les 256 états possibles de l'octet sont distribués avec une probabilité parfaitement égale.

V.1.2 3.2 Interprétation Critique des Seuils d'Entropie

L'investigation repose sur la déviation par rapport à la norme. Les fichiers courants (texte, exécutable) possèdent une structure statistique inhérente (ex: les instructions CPU ou les lettres fréquentes) qui maintient leur entropie **faible** ($H < 5$ bits/octet). L'augmentation de l'entropie est le signal d'un processus intentionnel.

Table 1: Classification des Données par Seuil d'Entropie

Plage d'Entropie (H)	**Nature Probable des Données**	**Implication en Investigation**
$H < 5$ bits/octet	Données textuelles, bases de données, exécutables (ELF, PE)	Normalité. Distribution très inégale (biaisée).
$5 \leq H < 7$ bits/octet	Fichiers multimédias compressés (JPEG, MP3), archives ZIP non chiffrées	Compression. Concentration sur l'en-tête et les métadonnées.
$H \approx 7.5$ à 8 bits/octet	Données chiffrées, malwares packés, stéganographie	**Alerte Rouge (Forte Aléa)**. Indique l'application d'une transformation intentionnelle.

V.1.3 3.3 Analyse Développée du Seuil Critique H ≈ 7.8 bits/octet

Une valeur d'entropie dans cette plage est considérée comme le **marqueur statistique de l'aléatoire parfait** (ou pseudo-aléatoire) et impose une investigation immédiate. Cette haute entropie signale trois phénomènes critiques :

V.1.3.1 1. Détection de Chiffrement (Ransomware et Sécurité)

Les algorithmes cryptographiques modernes (ex: AES) sont mathématiquement conçus pour produire un texte chiffré dont la distribution des octets est **indiscernable d'un bruit blanc**, c'est-à-dire une entropie maximale.

- **Diagnostic Post-Incident** : Si l'investigateur constate que les fichiers d'une victime (documents, photos) sont passés d'une entropie typique de 4.5 à 7.8, c'est une preuve factuelle d'une **action de chiffrement malveillante**, identifiant immédiatement l'impact d'un **ransomware**.
- **Preuve Explicative** : Cela démontre que l'investigation est passée du mode *lecture* (lecture de données non-chiffrées) au mode *cryptanalyse* ou *récupération de clés*.

V.1.3.2 2. Identification de Malwares Packés et Obfusqués

Les acteurs malveillants utilisent des outils de **packing** pour compresser et/ou chiffrer les exécutables afin d'échapper à la détection par signatures des antivirus.

- **Analyse en Mémoire** : Une analyse d'entropie locale sur différentes sections d'un binaire peut révéler qu'une section de code est fortement compressée/chiffrée ($H > 7$), tandis que les en-têtes restent à faible entropie. Cette anomalie est le signal d'une **charge utile chiffrée** qui sera décompressée uniquement en mémoire au moment de l'exécution (technique du *run-time decryption*).

V.1.3.3 3. Détection de Stéganographie Avancée

La stéganographie efficace (dissimulation d'informations) vise à minimiser la trace statistique de l'information cachée.

- **Analyse d'Entropie Locale** : L'insertion d'un message *chiffré* (haute entropie) dans un fichier porteur (ex: une image JPEG) peut provoquer des **pics localisés** d'entropie, en particulier dans les bits de poids faible. La cartographie spatiale de l'entropie sur l'ensemble des octets du fichier devient alors une technique de détection clé.

V.1.4 3.4 Conclusion : De la Technique à l'Épistémologie

L'entropie de Shannon incarne la rencontre entre les **mathématiques de l'information** et la **construction de la vérité légale**. Elle ne fournit pas la preuve de l'identité de l'attaquant, mais elle fournit la **preuve du processus d'attaque** (chiffrement ou obfuscation).

En tant qu'outil de diagnostic non-intrusif, il permet à l'investigateur de **prioriser** ses ressources et de définir la **stratégie analytique** (cryptanalyse vs. analyse de code) la plus appropriée, illustrant parfaitement la rigueur méthodologique exigée par l'investigation numérique.