Sécurité Informatique – IFT3275

Travail remis à Alain Tapp

DEVOIR 1

Sarah Bedn – 20214949

Khalil Rerhrhaye – 20179868

Le 10 novembre 2024

Session Automne 2024

Question 2

Pour la question 2, le code utilise une méthode de déchiffrement basée sur l'analyse de fréquence, une technique courante pour décrypter les messages chiffrés par substitution. L'idée principale est d'analyser la fréquence des éléments du cryptogramme et de les comparer aux fréquences des caractères d'un texte de référence, appelé corpus. En établissant des correspondances de fréquence entre le texte chiffré et le corpus, on peut recréer une partie du message original.

La première fonction, analyse_frequence_texte, calcule la fréquence des caractères individuels et des paires de caractères (appelées bicaractères) dans un texte donné. Cette fonction utilise Counter pour compter les occurrences de chaque caractère et bicaractère, puis elle calcule le pourcentage de chaque occurrence par rapport au total des caractères ou bicaractères du texte. Les résultats sont combinés dans un dictionnaire pour fournir une vue complète de la fréquence des caractères et des paires, puis triés par ordre décroissant pour que les éléments les plus courants soient en tête.

La fonction suivante, analyse_frequence_octets, analyse le cryptogramme en le divisant en blocs de 8 bits, appelés "octets." Elle compte la fréquence de chaque octet et calcule le pourcentage de chaque occurrence par rapport au total des octets. Ensuite, les octets sont triés par fréquence pour identifier les éléments les plus fréquents du cryptogramme, ce qui permet de les comparer aux caractères les plus courants du corpus.

La fonction principale, decrypt, commence par charger deux textes à partir d'URLs. Ces textes, issus d'un corpus en ligne, servent de référence pour l'analyse de fréquence. Si l'un des textes ne peut pas être chargé, une erreur est renvoyée. Ensuite, decrypt combine les deux textes pour créer un corpus complet, puis effectue une analyse de fréquence sur ce corpus à l'aide de analyse_frequence_texte. Une analyse de fréquence est également appliquée au cryptogramme grâce à analyse_frequence_octets. À partir de ces deux analyses, une correspondance est établie entre les octets fréquents du cryptogramme et les caractères fréquents du corpus.

Enfin, cette correspondance est utilisée pour traduire le cryptogramme en texte lisible : chaque octet du cryptogramme est remplacé par son caractère correspondant du corpus. Si un octet ne correspond à aucun caractère du corpus, un point d'interrogation (?) est utilisé comme substitut. Le message déchiffré est alors restitué.