

Guide complet de cryptanalyse pour textes français inconnus

Introduction

La cryptanalyse est l'art de déchiffrer des messages codés sans connaître préalablement la méthode de chiffrement ou la clé utilisée. Ce document présente une approche méthodique pour identifier et déchiffrer des textes en français dont on ignore le système de codage. Destiné aux passionnés de cryptographie, aux participants de jeux d'énigmes ou aux animateurs de camps routiers, ce guide fournit des outils pratiques et des techniques éprouvées pour résoudre efficacement ce type de challenge.

1. Première approche : Identifier le type de chiffrement

1.1 Observation visuelle du texte chiffré

La première étape consiste à observer attentivement le message pour déterminer sa nature :

Type d'éléments observés	Type de chiffrement probable
Uniquement des lettres de l'alphabet	Substitution monoalphabétique, César, Vigenère, transposition
Présence de chiffres uniquement	Chiffrement numérique, code du livre, coordonnées
Mélange de lettres et chiffres	Substitution alphanumérique, chiffre ADFGVX
Symboles non alphanumériques	Code symbole (morse, pigpen, semaphore)
Texte d'apparence normale	Stéganographie, message caché

1.2 Analyse statistique préliminaire

Avant toute tentative de déchiffrement :

- Comptez le nombre total de caractères
- Identifiez le nombre de symboles différents utilisés
- Notez la fréquence d'apparition de chaque symbole
- Repérez les motifs récurrents (groupes de caractères qui se répètent)

1.3 Tableau d'identification du chiffrement

Caractéristiques	Type probable	Méthode d'approche
26 symboles ou moins avec distribution inégale	Substitution monoalphabétique	Analyse fréquentielle
Peu de symboles (5 à 8)	Polybe, ADFGVX, Bacon	Tableau de correspondance
Texte avec répétitions régulières	Vigenère	Recherche de la longueur de clé (test de Kasiski)
Distribution uniforme des lettres	Transposition	Réarrangement des lettres
Séquences de points et traits	Morse	Décodage standard
Mots normaux mais message absurde	Texte à double sens ou grille	Extraction sélective

2. Analyse fréquentielle approfondie pour le français

L'analyse fréquentielle est l'outil le plus puissant pour déchiffrer les substitutions. Elle repose sur la fréquence d'apparition des lettres dans une langue donnée.

2.1 Distribution des lettres en français

Voici la fréquence des lettres en français (en pourcentage) :

E (14,7%) - S (7,9%) - A (7,6%) - I (7,5%) - N (7,1%) - T (7,0%) - R (6,6%)
U (6,0%) - L (5,5%) - O (5,4%) - D (3,7%) - C (3,3%) - P (3,0%) - M (2,6%)
V (1,6%) - H (1,1%) - G (1,0%) - F (1,0%) - B (0,9%) - J (0,7%) - Q (0,7%)
X (0,4%) - Z (0,3%) - K (0,1%) - W (0,1%) - Y (0,3%)

2.2 Bigrammes (paires de lettres) les plus fréquents en français

ES (3,0%) - DE (2,7%) - LE (2,4%) - EN (2,2%) - NT (2,0%) - RE (1,9%)
ON (1,9%) - ER (1,8%) - AN (1,7%) - OU (1,5%) - QU (1,5%) - IT (1,3%)

2.3 Trigrammes (triplets de lettres) les plus fréquents

ENT (1,0%) - LES (0,8%) - QUE (0,7%) - ANT (0,5%) - ION (0,5%)

2.4 Mots courts fréquents en français

LE, LA, LES, DE, DES, UN, UNE, ET, EST, EN, À, AU, AUX, CE, CES, IL, ELLE

2.5 Méthode pratique d'analyse fréquentielle

1. Comptez l'occurrence de chaque symbole dans le texte chiffré
2. Établissez un classement des symboles par fréquence décroissante

3. Comparez avec le classement des lettres en français
4. Commencez par remplacer les symboles les plus fréquents par les lettres correspondantes probables (E, S, A, I, N, T, R)
5. Identifiez les mots courts partiellement déchiffrés
6. Formulez des hypothèses et testez-les progressivement

3. Techniques de déchiffrement par type de chiffre

3.1 Chiffre de substitution monoalphabétique

Méthode systématique

1. Réalisez l'analyse fréquentielle complète
2. Identifiez les motifs correspondant à des mots courts courants (LE, LA, ET, DE)
3. Repérez les lettres isolées (probablement A, Y, À)
4. Identifiez les doubles lettres (LL, SS, MM, TT, FF)
5. Complétez progressivement les mots partiellement déchiffrés

Astuces spécifiques

- Les lettres qui apparaissent en fin de mot sont souvent E, S, T, X
- Les apostrophes sont généralement suivies par U, E, A, O
- Après Q, on trouve presque toujours U
- Les suites de 3 consonnes sont rares en français

3.2 Chiffre de César

Une variante spécifique de la substitution monoalphabétique où l'alphabet est simplement décalé.

Méthode de force brute

1. Essayez systématiquement les 25 décalages possibles
2. Pour chaque décalage, évaluez si le texte obtenu ressemble à du français

Astuce mathématique

Si vous soupçonnez un chiffre de César, calculez l'écart entre la lettre la plus fréquente du texte chiffré et E. Cet écart correspond probablement au décalage.

3.3 Chiffre de Vigenère

Ce chiffrement utilise plusieurs alphabets de substitution selon une clé.

Méthode de Kasiski

1. Repérez les séquences de lettres qui se répètent dans le texte chiffré
2. Calculez les distances entre ces répétitions
3. Trouvez le PGCD de ces distances pour estimer la longueur de la clé
4. Divisez le texte en colonnes selon cette longueur
5. Appliquez l'analyse fréquentielle sur chaque colonne séparément

Méthode de l'indice de coïncidence

1. Divisez le texte en groupes de longueur n (en testant différentes valeurs)
2. Calculez l'indice de coïncidence pour chaque longueur testée
3. La longueur qui donne un indice proche de 0,074 (valeur typique du français) est probablement la bonne
4. Procédez ensuite comme avec la méthode de Kasiski

3.4 Chiffres de transposition

Méthodes d'approche

1. Essayez différentes dispositions rectangulaires du texte
2. Testez des lectures en colonnes, en diagonales, en spirale
3. Si vous suspectez une clé, réarrangez les colonnes selon l'ordre alphabétique des lettres de la clé

Indices révélateurs

- La fréquence des lettres reste identique à celle du français
- Les bigrammes et trigrammes habituels sont absents
- La longueur du message est souvent un indice (produit de deux nombres)

3.5 Codes à symboles (Morse, Sémaphore, Pigpen)

Approche systématique

1. Identifiez le nombre de symboles différents utilisés
2. Si 2 symboles : probablement du morse ou du binaire
3. Si 9 symboles : probablement du pigpen
4. Groupez les symboles selon la logique apparente

Tableau des principaux codes à symboles

Nombre de symboles	Code probable	Caractéristiques
2	Morse	Points et traits ou 0 et 1
4	Chiffre carré	Symboles formés de L
9	Pigpen	Symboles formés de carrés et points
8	Sémaphore	Positions de drapeaux
26	Symboles de substitution	Un symbole par lettre

4. Déchiffrement des codes plus complexes

4.1 Chiffre ADFGVX

Ce chiffre combine substitution et transposition, utilisant uniquement les lettres A, D, F, G, V et X.

Méthode

1. Reconstituez une grille 6×6 avec les lettres et chiffres
2. Identifiez la clé de transposition par analyse des fréquences relatives
3. Réorganisez les colonnes
4. Traduisez les paires de lettres en utilisant la grille

4.2 Code du livre

Indices révélateurs

- Séries de nombres généralement groupés par 3 (page, ligne, position)
- Références répétées à certains nombres (pages populaires)

Approche

1. Identifiez le livre de référence (indices contextuels)
2. Si impossible, testez des livres classiques ou dictionnaires standards
3. Vérifiez la cohérence des numéros de page avec le livre supposé

4.3 Stéganographie

Méthodes de détection

1. Examinez la première lettre de chaque mot ou phrase (acrostiche)
2. Vérifiez si certaines lettres se démarquent (taille, police, espacement)
3. Testez l'extraction de texte selon un motif régulier (une lettre sur trois, etc.)
4. Cherchez des mots cachés en diagonale ou verticalement

5. Outils pratiques pour l'analyse manuelle

5.1 Tableau de fréquences

Créez un tableau sur papier avec :

- Première ligne : les symboles du texte chiffré
- Deuxième ligne : le nombre d'occurrences
- Troisième ligne : le pourcentage du total
- Quatrième ligne : les lettres françaises probables

5.2 Grille de décodage progressive

Utilisez une disposition en plusieurs lignes :

1. Le texte chiffré original
2. Une ligne pour noter vos hypothèses de déchiffrement
3. Une ligne pour marquer la fiabilité de vos hypothèses (sûr/probable/incertain)

5.3 Méthode des "mots ponts"

1. Identifiez les mots partiellement déchiffrés
2. Listez toutes les possibilités pour ces mots
3. Cherchez des mots qui partagent certaines lettres déjà identifiées
4. Utilisez ces "ponts" pour étendre progressivement votre déchiffrement

6. Approche méthodique face à un texte inconnu

6.1 Protocole d'analyse en 12 étapes

1. **Observation visuelle** : notez les particularités du texte (ponctuation, groupements)
2. **Comptage basique** : longueur du texte, nombre de symboles distincts
3. **Analyse fréquentielle** : établissez le classement des symboles
4. **Identification probable** : déterminez le type de chiffrement le plus probable
5. **Formation d'hypothèses** : proposez des correspondances pour les symboles les plus fréquents
6. **Recherche de mots courts** : identifiez les structures correspondant à LE, LA, ET, DE, etc.
7. **Vérification de cohérence** : assurez-vous que vos hypothèses sont compatibles
8. **Extension progressive** : complétez petit à petit les mots partiellement déchiffrés
9. **Ajustement** : réviser vos hypothèses si nécessaire
10. **Confirmation** : vérifiez que le texte déchiffré a du sens
11. **Identification de la méthode** : déterminez précisément le système utilisé
12. **Reconstruction de la clé** : si applicable, retrouvez la clé complète

6.2 Arbre de décision pour l'identification

1. Le texte utilise-t-il uniquement des lettres standards ?
 - └─ OUI → Passez à la question 2
 - └─ NON → Passez à la question 7
2. La distribution des lettres est-elle proche du français ?
 - └─ OUI → Probable transposition → Testez des arrangements rectangulaires
 - └─ NON → Passez à la question 3
3. Y a-t-il des motifs qui se répètent à intervalle régulier ?
 - └─ OUI → Probable Vigenère → Appliquez méthode de Kasiski
 - └─ NON → Passez à la question 4
4. Le texte utilise-t-il exactement 26 symboles différents ?
 - └─ OUI → Probable substitution monoalphabétique → Analyse fréquentielle
 - └─ NON → Passez à la question 5
5. Le texte utilise-t-il moins de 10 symboles différents ?
 - └─ OUI → Probable code à symboles réduits → Voir section 3.5
 - └─ NON → Passez à la question 6
6. Le texte semble-t-il avoir un sens mais contient un message caché ?
 - └─ OUI → Probable stéganographie → Voir section 4.3
 - └─ NON → Combinez plusieurs approches (chiffre hybride)
7. Le texte utilise-t-il uniquement des chiffres ?
 - └─ OUI → Probable code numérique ou du livre → Voir section 4.2
 - └─ NON → Passez à la question 8
8. Le texte utilise-t-il un ensemble restreint de symboles (ADFGVX) ?
 - └─ OUI → Probable chiffre ADFGVX → Voir section 4.1
 - └─ NON → Probable code à symboles étendu → Voir section 3.5

7. Études de cas pratiques

7.1 Exemple 1 : Déchiffrement d'une substitution monoalphabétique

Texte chiffré :

PH EBLYXHZLCYQVH HMX PC MBVHGBH FZV HXZFEVH PHM XHAEH BQFHMH HX PHM DYXQAFHM FH BHZLNCA SHEX
ZAVCGGCVL C P'CVFH FH DEXVPPCKHM RMVUZH. BAXVXH FH XCVGMVCXVAG HA PCGKCKHM MAGLHXM.

Analyse pas à pas :

1. Identification des fréquences : H (15,8%), F (7,9%), C (7,3%), X (7,3%), Z (5,9%), V (5,9%)

2. Correspondance probable : $H \rightarrow E$, $F \rightarrow D$, $C \rightarrow A$, $X \rightarrow T$, $Z \rightarrow U$, $V \rightarrow I$

3. Repérage de patterns : $PHM \rightarrow LES$, $FH \rightarrow DE$

4. Construction progressive du texte

Texte déchiffré :

LA CRYPTOGRAPHIE EST LA SCIENCE QUI ETUDIE LES TEXTE CODEES ET LES METHODES DE CHIFFREMENT A L'AIDE DE MODELES MATHEMATIQUES. NOTION DE TRADUCTION EN LANGAGES SECRETS.

7.2 Exemple 2 : Déchiffrement d'un Vigenère

Texte chiffré :

MBQCS ECRNM ZGIVS QAHKF GFRDV IQVZS GZUOV SGNZM XWPZF VSHCZ SBWTM

Analyse pas à pas :

1. Recherche de motifs répétés : MBQCS (au début) et ZGIVS
2. Application du test de Kasiski : estimation de la clé à 5 lettres
3. Division en 5 colonnes et analyse fréquentielle de chaque colonne
4. Détermination de la clé : CHAMP
5. Déchiffrement lettre par lettre

Texte déchiffré :

VOICI UNMES SAGEC HIFFR EPARL AMETH ODEDE VIGEN ERELA CLEDU CHIFF REMEN TEST

Après espacement correct :

VOICI UN MESSAGE CHIFFRE PAR LA METHODE DE VIGENERE LA CLE DU CHIFFREMENT EST

8. Conseils pour les cas difficiles

8.1 Quand aucune méthode ne fonctionne

1. **Combinez les approches** : Le texte peut utiliser plusieurs couches de chiffrement
2. **Cherchez un indice contextuel** : Le support du message, sa provenance ou son contexte peuvent donner des indices
3. **Testez des variantes linguistiques** : Le texte est peut-être en ancien français, argot ou dialecte
4. **Cherchez des erreurs** : Le texte peut contenir des fautes ou des lettres manquantes

8.2 Indices subtils à repérer

- **Motifs inhabituels de ponctuation** : Peuvent cacher un code morse
- **Espacements irréguliers** : Peuvent indiquer un code binaire
- **Variations typographiques minimales** : Peuvent être intentionnelles
- **Présence d'éléments graphiques** : Peuvent faire partie du code

8.3 Traitement des textes très courts

L'analyse fréquentielle est moins fiable sur les textes courts (moins de 100 caractères).

Méthodes alternatives :

1. Essayez la force brute (toutes les combinaisons possibles)
2. Utilisez le contexte pour deviner des mots probables
3. Testez des mots-clés en rapport avec la thématique

9. Outils numériques d'assistance

Même si l'approche manuelle est privilégiée, certains outils peuvent aider :

9.1 Applications utiles hors ligne

- **CryptoCrack** : Analyse automatique de substitutions
- **CrypTool** : Suite complète d'outils cryptographiques
- **dCode** (version hors ligne) : Multiples décodeurs

9.2 Création d'outils simples avec papier et crayon

- **Disque de César** : Deux disques concentriques avec alphabets pour tester les décalages
- **Tableau de Vigenère** : Grille 26×26 pour déchiffrement manuel
- **Grille de fréquences** : Tableau pré-imprimé pour comptage rapide

10. Exercices pratiques progressifs

Pour maîtriser ces techniques, entraînez-vous avec ces exercices de difficulté croissante :

10.1 Niveau débutant

1. Déchiffrez un message en César avec indice
2. Identifiez le type de chiffrement parmi 3 exemples donnés
3. Complétez un message partiellement déchiffré

10.2 Niveau intermédiaire

1. Déchiffrez une substitution monoalphabétique sans indice
2. Trouvez la clé d'un Vigenère court
3. Résolvez un message en transposition simple

10.3 Niveau avancé

1. Déchiffrez un texte combinant deux méthodes
2. Résolvez un message dont certaines parties manquent
3. Identifiez et déchiffrez un texte utilisant un code rare

Conclusion

La cryptanalyse de textes français inconnus est un art qui combine rigueur méthodique et intuition. Ce guide vous a fourni les outils essentiels pour aborder systématiquement ce défi passionnant. Avec de la pratique et de la persévérance, vous pourrez déchiffrer la plupart des textes codés traditionnels.

Rappelez-vous que chaque texte chiffré est unique et peut nécessiter une combinaison personnalisée de techniques. La patience et la méthodologie sont vos meilleures alliées dans cette quête de dévoilement des messages secrets.

Annexes

A. Tableau complet des fréquences en français

[Tableau détaillé des fréquences de lettres, bigrammes et trigrammes]

B. Fiches techniques détachables

[Fiches synthétiques pour analyse rapide sur le terrain]

C. Bibliographie commentée

- Simon Singh, "Histoire des codes secrets", 1999 - Excellente introduction historique
- David Kahn, "The Codebreakers", 1996 - La référence complète en cryptologie
- Gilles Dowek, "Les métamorphoses du calcul", 2007 - Pour comprendre les fondements mathématiques
- Pierre Berloquin, "Codes et langages secrets", 2015 - Approche ludique et pratique