CS4021D NUMBER THEORY AND CRYPTOGRAPHY

ASSIGNMENT 1

SAGEMATH IMPLEMENTATION OF BASIC CRYPTOGRAPHIC SYSTEMS AND IT'S CRYPTANALYSIS

SINGAM SAI BALA SUBRAHMANYAM
B180522CS
S5 BTECH CSE A

1. Affine Cipher

Logic:

$$C = (P \times k_1 + k_2) \bmod 26 \qquad P = ((C - k_2) \times k_1^{-1}) \bmod 26$$

where k_1^{-1} is the multiplicative inverse of k_1 and $-k_2$ is the additive inverse of k_2

Cryptanalysis Techniques used

- Brute Force

Key domain is 26*12 = 312 values, we can brute force through
this key domain

- Statistical attack

Affine cipher is a mono alphabetic substitution cipher that preserves the frequencies of letters. Hence we can replace the most frequent letter with that of top 5 frequent letters as per english alphabet and can find possible key pairs by finding k1 and k2 from equation
C=(P*k1 +k2)mod26

- Known Plain Text

If we know two plain text and corresponding cipher text letters we can form 2 equations

(P1 * K1 + K2)mod26 = C1
(P2 * K1 + K2)mod26 = C2

And P_matrix = [P1 1
P2 1]

C_matrix = [C1
C2]

We can get key as key = P_matrix.inverse * C_matrix

Input file format:

First line : Plain text
Second line: key pair as key1,key2

Output file format:

Prints Encrypted, Decrypted and text with spaces on separate lines

2. Hill Cipher

Logic

$$
\begin{aligned}
C_1 &= P_1 k_{11} + P_2 k_{21} + \cdots + P_m k_{m1} \\
C_2 &= P_1 k_{12} + P_2 k_{22} + \cdots + P_m k_{m2} \\
&\cdots \\
C_m &= P_1 k_{1m} + P_2 k_{2m} + \cdots + P_m k_{mm}
\end{aligned}
$$

C = (P * key)mod26
P = (C* key.inverse)mod 26

Cryptanalysis Techniques used

- Chosen plain text

If we know value of block size m and m blocks of plain text and cipher text of same or different messages, we can form key using logic

Key = P_matrix.inverse * C_matrix

Input file format

First line : plain text
Second line: block size m
Third line: matrix elements separated by comma (,)

Output file format

Prints Encrypted, Decrypted and text with spaces on separate lines

3. Shift Cipher

Logic

C = (P + k)mod26
P = (C-k)mod26

Cryptanalysis Techniques used

- Brute Force

Key domain is 26 values, we can brute force through this key domain

- Statistical attack

Shift cipher is a mono alphabetic substitution cipher that preserves the frequencies of letters. Hence we can replace the most frequent letter with that of top 5 frequent letters as per english alphabet and can find possible key by finding k from equation

$$C=(P+K)\mod 26$$

- Digram attack

Shift cipher preserves the frequencies of diagrams hence we can find key by analysing the digram frequencies

Input file format

First line: Plain text
Second line: key

Output file format

Prints Encrypted, Decrypted and text with spaces on separate lines

4. Substitution Cipher

Logic: After agreeing on a certain key, a mapping between each alphabet and key is created

Cryptanalysis Techniques used

- Brute Force

  It takes 26! Values to verify but its very computationally expensive

- Statistical attack

  Substitution cipher is a mono alphabetic substitution cipher that preserves the frequencies of letters. Hence we can do frequency analysis and try to find the key mappings

- Known Plain Text

  If "ABCDEFGHIJKLMNOPQRSTUVWXYZ" plain text can be encrypted it returns the key used as cipher text.

Input file format

First line: Plain text
Second line: key

Output file format

Prints Encrypted, Decrypted and text with spaces on separate lines

5. Transposition Cipher

Logic: Transposition does not change the characters but instead change the positions of characters

Keyless: Agree on number of columns, write row by row create encrypted key by reading column by column decryption exact reverse

Keyed: Create blocks of plain text and permute in each block

Cryptanalysis Techniques used

- Brute force for keyless

  Brute force by taking number of columns from 1 - 26

- Brute force for keyed

  Brute force by taking permutations in size of factors of length of plaintext like 1! 2! 3! ..... P! Where P is size of plaintext
  Only upto 7! Is considered since its costly

Input file format

First line: Plain text
Second line: column size for keyless
Third line: Block size for keyed
Fourth line: Permutation key as numbers separated by coma(,)

Output file format

Prints Encrypted, Decrypted and text with spaces of keyless and keyed on separate lines

6. Vigenere Cipher

Logic:

Where K is any string given as input for key

Cryptanalysis Techniques used

- Kasiski test

We search for repeated text segments, of at least three characters, in the ciphertext. Suppose that two of these segments are found and the distance between them is d. It can be assumed that d|m where m is the key length. If more repeated segments can be found we take gcd of those distances d1, d2, d3 .... If gcd is m then key size is multiple of m and we divide cipher text to m parts and perform frequency analysis on each of those segments. Only key size upto 5 is considered in my implementation of kasiski test since its very costly

Input file format

First line: Plain text
Second line: key

Output file format

Prints Encrypted, Decrypted and text with spaces on separate lines