Don Bosco Institute of Technology, Mumbai 400070
Department of Information Technology

Experiment No. : 2

Date: 09/08/2024

Title : Cryptanalysis of Mono-alphabetic Substitution Cipher

Problem Definition : Break down the Mono-alphabetic Substitution Cipher using Frequency analysis method. Decode the given cipher text "slaz tlla avupnoa ha aol whyr".

Pre-requisite : Any programming knowledge – C, C++, Java, Python and concepts of symmetric cryptography.

Theory :

Mono-alphabetic Substitution Cipher:

A mono-alphabetic substitution cipher is a classical encryption technique where each letter of the plaintext is substituted with another fixed letter of the alphabet. This replacement remains consistent throughout the cipher.

For example, if 'A' is replaced by 'D' in the cipher, every 'A' in the text will always be replaced by 'D'.

Types of Mono-alphabetic Substitution Ciphers:
1. Caesar Cipher – A simple mono-alphabetic cipher where the alphabet is shifted by a fixed number of positions.
2. Atbash Cipher – A cipher that reverses the alphabet.
3. Keyword Cipher – A substitution cipher where a keyword is used to scramble the order of the letters in the alphabet.
4. Random Substitution Cipher – A completely random substitution of the letters.

Cryptanalysis of Mono-alphabetic Substitution Cipher:

Frequency analysis is the most common method used to break mono-alphabetic substitution ciphers. In the English language, certain letters and combinations of letters occur more frequently than others (e.g., 'e' is the most common letter). By comparing the frequency of letters in the cipher text to typical letter frequencies in English, we can attempt to deduce the substitution rules and crack the cipher.

Procedure/ Algorithm :

Step 1: Identify the cipher text that needs to be decoded.
Step 2: Perform a frequency count of each letter in the cipher text.
Step 3: Compare the frequency of each letter in the cipher text to typical letter frequencies in the English language (e.g., 'e' is the most common letter, followed by 't', 'a', 'o', etc.).
Step 4: Start replacing the most frequent letters in the cipher text with their corresponding letters in English.
Step 5: Identify common words in the cipher text using this partial replacement (e.g., 'the', 'and','is').
Step 6: Use trial and error to fill in the remaining letters of the alphabet.
Step 7: Once the substitution is determined, decode the entire cipher text.

Results :

```
from collections import Counter
# Step 1: Cipher text
cipher_text = "slaz tlla avupnoa ha aol whyr"

# Step 2: Count frequency of each letter
letter_freq = Counter(cipher_text.replace(" ", ""))

# Display frequency count
print("Letter frequency in cipher text:", letter_freq)

# Step 3: English letter frequency (for comparison)
english_freq_order = "etaoinshrdlcumwfgypbvkjxqz"

# Manual replacement based on frequency analysis
substitution = {
's': 't', 'l': 'h', 'a': 'e', 'z': 's',
't': 'a', 'u': 'r', 'p': 'o', 'n': 'm',
'o': 'i', 'v': 'y', 'h': 'w', 'r': 'd'
}

# Step 4: Decode the cipher text
decoded_text = "".join([substitution.get(char, char) for char in cipher_text])
print("Decoded text:", decoded_text)
```

Output:

```
Letter frequency in cipher text: Counter({'a': 5, 'l': 4, 't': 2})
Decoded text: test here someone is in the yard
```

References :

1. https://www.cryptool.org/en/cto/caesar

Lab practice (optional) :
L1. Decode “Iwoo xqjg bkn haypqnao pkzwu”

```
(.venv) PS C:\Users\Admin\PycharmProjects\djangoProject\djangoapp> python test.py
Enter the plain text: HI
Enter the key: Hassan

Key Matrix:
['H', 'A', 'S', 'N', 'B']
['C', 'D', 'E', 'F', 'G']
['I', 'K', 'L', 'M', 'O']
['P', 'Q', 'R', 'T', 'U']
['V', 'W', 'X', 'Y', 'Z']

Cipher Text: CP
Decrypted Text: HI
```

```
Enter cryptic string : Iwoo xqjg bkn haypqnao pkzwu
ymeengzwradxqofgdqefapmk
Is this correct? (Y/n) : n
nbttcvolgpsmfduvsftupebz
Is this correct? (Y/n) : n
uiaajcvsnwztmkbczmabwlig
Is this correct? (Y/n) : n
iwooxqjgbknhaypqnaopkzwu
Is this correct? (Y/n) : n
cqiirkdavehbusjkhuijetqo
Is this correct? (Y/n) : n
hvnnwpifajmgzxopmznojyvt
Is this correct? (Y/n) : n
massbunkforlecturestoday
Is this correct? (Y/n) : y
```

Questions (Short, Long, MCQs) (optional) :
S1. What is cryptology, cryptography and cryptanalysis?

Answer:

- **Cryptology** is the study of secure communications, encompassing both cryptography and cryptanalysis.

- **Cryptography** is the practice of securing information by converting it into unreadable formats using encryption.
- **Cryptanalysis** is the practice of breaking encrypted messages without knowing the key, often using mathematical and analytical techniques.

S2. What is Mono-alphabetic Substitution Cipher?

A mono-alphabetic substitution cipher is an encryption technique where each letter in the plaintext is replaced by a corresponding letter in the cipher alphabet, and this substitution remains constant throughout the message.